Titre: Title:	A Class of Networked Multi-Agent Control Systems: Interference Induced Games, Filtering, Nash Equilibria
Auteur: Author:	Mehdi Abedinpour Fallah
Date:	2017
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Abedinpour Fallah, M. (2017). A Class of Networked Multi-Agent Control Systems: Interference Induced Games, Filtering, Nash Equilibria [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2523/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/2523/
Directeurs de recherche: Advisors:	Roland Malhamé
Programme: Program:	génie électrique

UNIVERSITÉ DE MONTRÉAL

A CLASS OF NETWORKED MULTI-AGENT CONTROL SYSTEMS : INTERFERENCE INDUCED GAMES, FILTERING, NASH EQUILIBRIA

MEHDI ABEDINPOUR FALLAH
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE ÉLECTRIQUE)
AVRIL 2017

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

A CLASS OF NETWORKED MULTI-AGENT CONTROL SYSTEMS : INTERFERENCE INDUCED GAMES, FILTERING, NASH EQUILIBRIA

présentée par : ABEDINPOUR FALLAH Mehdi
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

M. LE NY Jérôme, Ph. D., président
M. MALHAMÉ Roland P., Ph. D., membre et directeur de recherche
M. GOURDEAU Richard, Ph. D., membre
M. MAHAJAN Aditya, Ph. D., membre externe

DEDICATION

To my beloved family

ACKNOWLEDGEMENTS

It gives me great pleasure to thank those who contributed to the successful completion of my studies at Polytechnique Montréal. First and foremost, I owe my deepest gratitude to my Ph.D. supervisor, Professor Roland Malhamé, for his persistent support, encouragement, and patience during my doctoral studies. I enjoyed working with him, with all his considerate, thoughtful principles, and learning from him, with his vast knowledge about all aspects of my work. His brilliant supervision is behind all my achievements during my Ph.D. studies. I feel so blessed to work with him as his Ph.D. student. What I have learned from him in research, as well as in real life, is absolutely ineffable. I also feel fortunate to have had the opportunity to collaborate with Professors Francesco Martinelli and David Saussié during my Ph.D. research. I would like to thank them for their contributions in this thesis.

I would also like to express my sincere gratitude to my thesis committee members, Professors Richard Gourdeau, Aditya Mahajan and Jérôme Le Ny for their time and comments. I am also grateful to my Master's advisor, Professor Luis Rodrigues, for his constant encouragement.

I am grateful to Vilmar, Filippo, Sadegh, Sara, Stéphane, Nadir, Zoumana, Arman and Adham with whom I have shared my GERAD office and had many interesting conversations. I learned a lot from our endless talks and discussions. I acknowledge their support and encouragement and I wish them health and success. I am also thankful to Nathalie and Suzanne of secretary staff of Poly for their great work.

I am especially thankful to Shahrouz and Leila for always being there for me. I am also thankful to all my friends who turned my life in Montreal to an amazing and rewarding journey. In particular, I thank all my hiking, biking, rock 'n' roll and swing dancing friends. These refreshing activities gave me the energy to continue my research. I would also like to thank Levi, Tyler and Christopher for our enlightening discussions.

Last but not least, I would like to say *Thank you!* to my beloved mother and sister for their affection and support during my life.

RÉSUMÉ

Nous considérons une classe de systèmes de contrôle stochastiques linéaires scalaires en réseau dans lesquels un grand nombre d'agents contrôlés envoient leurs états à un concentrateur central qui, à son tour, envoie des commandes de contrôle silencieuses basées sur ses observations et vise à minimiser un coût quadratique donné. La technologie de communication est l'accès multiple par répartition en code (CDMA) et, par conséquent, les signaux reçus sur le concentrateur central sont corrompus par des interférences. Les niveaux des signaux envoyés par les agents sont considérés proportionnels à leur état, et le traitement des signaux basés sur le CDMA réduit l'interférence d'autres agents d'un facteur de $1/N$ où N est le nombre d'agents. L'interférence existante crée par inadvertance une situation de jeu dans laquelle les actions d'un agent affectent son état et donc par interférence, la capacité d'autres agents à estimer les leurs, influençant à leur tour leur capacité à contrôler leur état. Ceci conduit à des problèmes d'estimation fortement couplés. Cela conduit également à une situation de contrôle dual puisque les contrôles individuels contrôlent l'état mais affectent également le potentiel d'estimation de cet état. La thèse comporte trois parties principales.

Dans la première partie, nous montrons que le fait d'ignorer le terme d'interférence et d'utiliser un principe de séparation pour le contrôle mène à des équilibres de Nash asymptotiques en N, pourvu que la dynamique individuelle soit stable ou "pas excessivement" instable. Que pour certaines classes de coût et de paramètres dynamiques, les lois de contrôle séparées optimales obtenues en ignorant le couplage interférentiel, sont asymptotiquement optimales lorsque le nombre d'agents passe à l'infini, formant ainsi pour un nombre de joueurs fini N, un équilibre ϵ-Nash. Plus généralement, les lois de contrôle séparées optimales peuvent ne pas être asymptotiquement optimales et peuvent en fait conduire à un comportement global instable. Nous considérons donc une classe de lois de contrôle décentralisées paramétrées selon lesquelles le gain séparé de Kalman est traité comme le gain arbitraire d'un observateur analogue à un observateur de Luenberger. Les régions de stabilité du système sont caractérisées et la nature des politiques optimales de contrôle coopératif au sein de la classe considérée est explorée.

La deuxième partie concerne l'extension du travail dans la première partie au-delà du seuil d'instabilité des contrôles coopératifs. Il est alors observé que les contrôles linéaires invari-

ants dans le temps basés sur les sorties des filtres de dimension croissante semblent toujours maintenir la stabilité du système et d'intrigantes propriétés sur les estimations des états sont observées numériquement. En particulier, nous abordons le cas d'un filtrage décentralisé exact sous une classe de contrôleurs basés sur un retour d'estimateur d'état avec gain invariant, et nous étudions numériquement à la fois la capacité de stabilisation et la performance de tels contrôleurs lorsque le gain de rétroaction de l'estimation de l'état est modifié. Alors que les filtres optimaux ont des besoins de mémoire qui deviennent infinis dans le temps, la capacité de stabilisation de leur approximation de mémoire finie est également testée.

La dernière partie porte sur le développement d'algorithmes basés sur des points fixes pour l'identification des stratégies de Nash. En particulier, pour un problème d'horizon fini, nous proposons un algorithme basé sur un point fixe qui tient compte d'une combinaison de coûts de contrôle et d'estimation pour calculer les équilibres de Nash symétriques, s'ils existent. Cela implique d'alterner à plusieurs reprises un *forward sweep* pour l'estimation d'état et un *backwards sweep* pour une optimisation basée sur la programmation dynamique.

ABSTRACT

We consider a class of networked linear scalar stochastic control systems whereby a large number of controlled agents send their states to a central hub, which in turn sends back noiseless control commands based on its observations, and aimed at minimizing a given quadratic cost. The communication technology is code division multiple access (CDMA), and as a result signals received at the central hub are corrupted by interference. The signals sent by agents are considered proportional to their state, and CDMA based signal processing reduces other agents' interference by a factor of $1/N$ where N is the number of agents. The existing interference inadvertently creates a game situation whereby the actions of one agent affect its state and thus through interference, the ability of other agents to estimate theirs, in turn influencing their ability to control their state. This leads to highly coupled estimation problems. It also leads to a dual control situation as individual controls both steer the state and affect the estimation potential of that state. The thesis is presented in three main parts.

In the first part, we show that ignoring the interference term and using a separation principle for control provably leads to Nash equilibria asymptotic in N, as long as individual dynamics are stable or "not exceedingly" unstable. In particular, we establish that for certain classes of cost and dynamic parameters, optimal separated control laws obtained by ignoring the interference coupling are asymptotically optimal when the number of agents goes to infinity, thus forming for finite N an ϵ-Nash equilibrium. More generally though, optimal separated control laws may not be asymptotically optimal, and can in fact result in unstable overall behavior. Thus we consider a class of parameterized decentralized control laws whereby the separated Kalman gain is treated as the arbitrary gain of a Luenberger like observer. System stability regions are characterized and the nature of optimal cooperative control policies within the considered class is explored.

The second part is concerned with the extension of the work in the first part past the instability threshold for the previous cooperative Luenberger like observers. It is observed that time invariant linear controls based on the outputs of growing dimension filters appear to always maintain system stability, and intriguing state estimate properties are numerically observed. More specifically, we tackle the case of exact decentralized filtering under a class of time invariant certainty equivalent feedback controllers, and numerically investigate both

stabilization ability and performance of such controllers as the state estimate feedback gain varies. While the optimum filters have memory requirements which become infinite over time, the stabilization ability of their finite memory approximation is also tested.

The final part focuses on developing fixed point based algorithms for identifying Nash strategies. In particular, for a finite horizon problem, we propose a fixed point based algorithm which accounts for a combination of control and estimation costs to compute symmetric Nash equilibria if they exist. It involves repeatedly alternating a forward sweep for state estimation and a backwards sweep for dynamic programming based optimization.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ARE Algebraic Riccati Equation
BS Base Station

CDMA Code Division Multiple Access

dB Decibel

FIR Finite Impulse Response

IIR Infinite Impulse Response

LQG Linear Quadratic Gaussian

LTA Long Time Average

MFG Mean Field Game

MPC Model Predictive Control

MSS Mean Square Stable

NCE Nash Certainty Equivalence

NCS Networked Control System

NE Nash Equilibrium

ODE Ordinary Differential Equation

OSNR Optical Signal to Noise Ratio

QNG Quadratic Network Game

SINR Signal to Interference plus Noise Ratio

$\mathbb{E}$ Expected value

$\mathcal{N}$ Normal distribution

Tr Trace of a matrix

$\mathbb{R}$ Set of real numbers

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

1.1 Background Information

Networked Control System (NCS) refers to a decentralized control system in which the components are connected through real-time communication channels or a data network. There may be a data link between the sensors (which collect information), the controllers (which make decisions), and the actuators (which execute the controller commands); and the sensors, the controllers, and the plant themselves could be geographically separated (Yüksel and Başar (2013)). A multi-agent system is a network of multiple interacting components (agents). Each agent is assumed to hold a state regarding one or more quantities of interest. Depending on the context, states may be referred to as values, positions, velocities, temperatures or etc. Many practical applications and examples of large population stochastic multi-agent systems arise in engineering, biological, social and economic fields, such as wireless sensor networks (Chong and Kumar (2003)), very large scale robotics (Reif and Wang (1999)), swarm and flocking phenomena in biological systems (Grönbaum and Okubo (1994); Passino (2002)), evacuation of large crowds in emergency situations (Helbing *et al.* (2000); Lachapelle (2010)), sharing and competing for resources on the Internet (Altman *et al.* (2006)), charging control for large populations of plug-in electric vehicles (Parise *et al.* (2014); Grammatico (2016)), and so on.

In conventional control systems, control laws are constructed based upon the overall states of the plants. However, in complex systems with many agents, each agent has a self-governed but limited capability of sensing, decision-making and communication. Therefore, in multi-agent decision making in the context of networked control systems with large number of agents an important issue is the development of decentralized solutions so that each individual agent may implement a strategy based on its local information together with statistical information on the population of agents. These are systems where different decision units (or equivalently decision makers or agents, which could be sensors, controllers, encoders, or decoders) are connected over a communication network, where information is decentralized. Just as stabilization and optimization are two fundamental issues for single-agent systems, for large population stochastic multi-agent systems we are also concerned with how to construct decentralized control laws that preserve closed-loop system stability while optimizing the performance of agents in a cooperative or non-cooperative context.

Game theory has emerged as a well-established discipline capable of providing a resourceful and effective framework for addressing control of large scale and distributed networked systems. Agent to agent interaction during competitive decision-making is usually due to the coupling in their dynamics or cost functions. Specifically, the dynamic coupling is used to specify an environment effect on the individual's decision-making as generated by the population of other agents. While each agent only receives a negligible influence from any other given individual, the overall effect of the population (i.e., the mass effect) is significant for each agent's strategy selection (see Huang *et al.* (2007) for example). In this thesis, we study a somewhat dual situation whereby large populations of partially observed stochastic agents, although a priori individually independent, are coupled via their observation structure. The latter involves an interference term depending on the empirical mean of all agent states. The study of such measurement-coupled systems is inspired by a variety of applications, including for instance the communications model for power control in CDMA technology based cellular telephone systems (Huang *et al.* (2004); Perreau and Anderson (2006)), where any conversation in a cell acts as interference on the other conversations in that cell. Indeed, despite the so-called signal processing gain achieved thanks to a user's specific coding advantage (and considered in our model to be of order $1/N$ where N is the total number of agents), the ability of the base station to correctly decode the signals sent by a given mobile, remains limited by interference formed by the superposition of all other in cell user signals. Viewed in this light, the studied problem can be considered as a game over a noisy channel.

1.2 Problem Formulation

Consider a model of a code division multiple access (CDMA) based communication and control system in the context of a large number of users with N users which share a channel and are assumed to be equally spaced on a circle around the base station as depicted in Figure 1.1, with a signal processing gain proportional to $1/N$. The base station itself sends the control signal to a collection of individual systems (users), hereon also referred to as agents. Downlink channels are considered noiseless, however the controlled individual systems are stochastic. The i^{th} mobile user of the network transmits a signal proportional to the (scalar) state $x_{k,i}$, that is to say, $\beta x_{k,i}$, where β is a constant parameter. Note that the transmitted power is proportional to $\beta^2 x_{k,i}^2$ and that the larger the state, the more energy will be involved in the transmission. The base station in turn computes the required control based on received signal which also is tainted by interference and noises. In particular, the

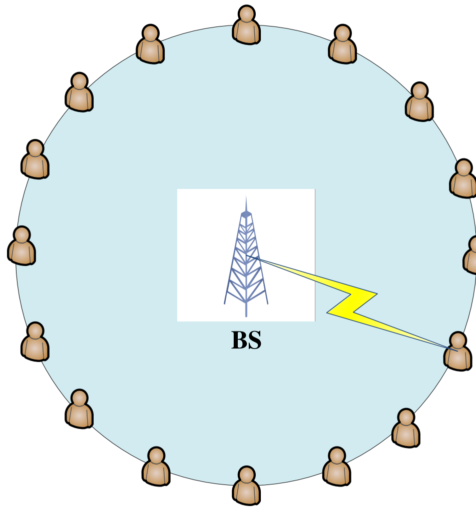

Figure 1.1 N users using CDMA technology

output signal corresponding to the i^{th} user (agent) is given by:

$$y_{k,i} = \alpha\beta x_{k,i} + \frac{h'}{N}\sum_{j\neq i}^{N}\alpha\beta x_{k,j} + v_{k,i}^{th} + v'_{k,i}, \tag{1.1}$$

where $\alpha > 0$ denotes the uplink channel gain of the network, $v_{k,i}^{th}$ is the background thermal noise process and $v'_{k,i}$ is the the local observation error after transmission ($v_{k,i}^{th}$ and $v'_{k,i}$ are modeled as zero mean Gaussian random variables) [see Abedinpour Fallah *et al.* (2016); Huang *et al.* (2004); Koskie and Gajic (2006); Perreau and Anderson (2006)]. Note that the resulting signal processing gain is assumed to be h'/N. Also, the actual controlling users are assumed to be independent and simply using the base station as a communication tool. Thus, by letting

$$c = \alpha\beta(1 - \frac{h'}{N}), \quad h = \alpha\beta h', \quad v_{k,i} = v_{k,i}^{th} + v'_{k,i}, \tag{1.2}$$

the physics of CDMA transmission viewed as a networked control system with N agents can be cast into a state space form with individual scalar dynamics described by

$$x_{k+1,i} = ax_{k,i} + bu_{k,i} + w_{k,i} \tag{1.3}$$

and partial scalar state observations given by:

$$y_{k,i} = cx_{k,i} + h\left(\frac{1}{N}\sum_{j=1}^{N}x_{k,j}\right) + v_{k,i} \tag{1.4}$$

for $k \geq 0$ and $1 \leq i \leq N$, where $x_{k,i}, u_{k,i}, y_{k,i} \in \mathbb{R}$ are the state, the control input and the measured output of the i^{th} agent, respectively. The random variables $w_{k,i} \sim \mathcal{N}(0, \sigma_w^2)$ and $v_{k,i} \sim \mathcal{N}(0, \sigma_v^2)$ represent independent Gaussian white noises at different times k and at different agents i. The Gaussian initial conditions $x_{0,i} \sim \mathcal{N}(\bar{x}_0, \sigma_0^2)$ are mutually independent and are also independent of $\{w_{k,i}, v_{k,i}, 1 \leq i \leq N,\ k \geq 0\}$. σ_w^2, σ_v^2 and σ_0^2 denote the variance of $w_{k,i}$, $v_{k,i}$ and $x_{0,i}$, respectively. Moreover, a is a scalar parameter and $b, c, h > 0$ are positive scalar parameters.

1.3 Objectives

The research presented in this thesis has been conducted with the objective of studying the decentralized control of partially observed multi-agent systems with mutually interfering measurements, from two distinct viewpoints: (i) *A game theoretic viewpoint* corresponding to the study of a problem dual to that of standard mean field games where agents' coupling occurs only through cost and dynamics. This leads to questions of existence of Nash or ϵ-Nash equilibria under arbitrary or particular classes of linear non anticipative decentralized output feedback laws; (ii) *A filtering/state estimation viewpoint* whereby the coupling of estimation with control leads implicitly to an embedded infinite sequence of estimation problems, in that each agent attempts to produce an estimate of what other agents' own state estimates are, and in turn that agent needs to estimate what other agents think its own state estimate is, etc. Thus in general, infinite memory filters are required, although the dynamics remain completely linear.

1.4 Contributions

The main contributions of the thesis are as follows.

1.4.1 Dualization of mean field game theory-based formulation from control to estimation situations

The decentralized control problem of partially observed multi-agent systems with N uniform agents described by linear stochastic dynamics, quadratic costs and partial linear observations involving the mean of all agents is formulated as a game.

1.4.2 Establishing stability and optimality properties of the separated policies

The conditions are established under which a Luenberger like observer of the form

$$\hat{x}_{k+1,i} = (a - bf)\hat{x}_{k,i} + K(y_{k+1,i} - c(a - bf)\hat{x}_{k,i}), \tag{1.5}$$

where $\hat{x}_{k,i}$ is an estimator of $x_{k,i}$ based only on local observations of the i^{th} agent and f, K are constant scalar gains, together with a constant state estimate feedback of the form

$$u_{k,i} = -f\hat{x}_{k,i}, \tag{1.6}$$

would be: (i) ideally optimal or a Nash equilibrium with respect to some long term average performance measure given by

$$J_i \triangleq \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2). \tag{1.7}$$

where $\overline{\lim}$ denotes the limit superior and $r > 0$ is a positive scalar parameter, (ii) at least stable. More specifically, the following questions are investigated thoroughly:

— *How long can agents remain indifferent to interference if numbers are sufficiently high?* Assuming $a > 0$ for the sake of analysis, the threshold a_{Nash} is the (unique) value of a if it exists, such that:

$$a_{Nash} - bf_{\sup}(a_{Nash}) = 0, \tag{1.8}$$

which represents the maximum value of a past which it is not always possible to apply so-called isolated optimal control policies without causing potential instability problems. More specifically, if $a < a_{Nash}$, we have that $(K^*(a), f^*(a,r))$ stabilizes the population for all r, where $K^*(a)$ is the Kalman gain as obtained when assuming zero interference in the local measurements (setting $h = 0$ in (1.4)) and $f^*(a,r)$ denotes the control gain obtained by the appropriate algebraic Riccati equation. For given a, $K^*(a)$, the stability region for f, if non empty, is an interval $(f_{\inf}(a), f_{\sup}(a))$ such that

$$f_{\inf}(a) = \frac{(a-1)(1-a(1-cK^*))}{b(1-a(1-cK^*)+hK^*)}\,, \quad if \quad a < 1; \tag{1.9}$$

$$f_{\inf}(a) = \frac{a-1}{b}\,, \quad if \quad a \geq 1; \tag{1.10}$$

$$f_{\sup}(a) = \frac{a+1}{b}, \quad if \quad a \leq -1; \tag{1.11}$$

$$f_{\sup}(a) = \frac{(a+1)(1+a(1-cK^*))}{b(1+a(1-cK^*)+hK^*)}, \; if \; a > -1. \tag{1.12}$$

Moreover, we have:

(i) for $a > a_{Nash}$, $a/b > f_{\sup}(a)$;

(ii) for $a < a_{Nash}$, $a/b < f_{\sup}(a)$;

(iii) a_{Nash} can assume any positive value by properly modifying h. In particular, $a_{Nash} > 1$ for all $h < (2 - cK^*(1))/K^*(1)$ and $a_{Nash} < 1$ for all $h > (2 - cK^*(1))/K^*(1))$.

— *What ranges of r can agents still remain indifferent to interference if numbers are sufficiently high, if $a > a_{Nash}$, and for what range of a?*

The answer is given as $r > r_{inf}$ and $a < a_{\sup}$. In particular, let $S(a)$ designate the (K, f) stability region associated with parameter a, also let $\bar{f}_{sup}(a) = \min\{a/b, f_{\sup}(a)\}$ and define

$$r_{\inf}(a) = \frac{-b(a - b\bar{f}_{sup}(a))}{\bar{f}_{sup}(a)[a(a - b\bar{f}_{sup}(a)) - 1]}. \tag{1.13}$$

Then there exists a threshold $a_{\sup} > \max\{a_{Nash}, 1\}$, which satisfies

$$a_{\sup}[a_{\sup} - bf_{\sup}(a_{\sup})] - 1 = 0, \tag{1.14}$$

such that, for all $a < a_{\sup}$, $(K^*(a), f^*(a,r)) \in S(a)$ for all $r > r_{\inf}(a) \geq 0$ (with $r_{\inf}(a) = 0$ for all $a \leq a_{Nash}$). Moreover, $r_{\inf}(a) \to +\infty$ as $a \to a_{\sup}$, and, if $a > a_{\sup}$, $(K^*(a), f^*(a,r)) \notin S(a)$ for all r. In essence, for $a_{Nash} < a < a_{\sup}$, $(K^*(a), f^*(a,r)) \in S(a)$ only if r is larger than a positive threshold $r_{\inf}(a)$, and if $a = a_{\sup}$, $r_{\inf}(a)$ reaches $+\infty$.

— *When does cooperation become crucial?*

Past $a_{\sup}$ and up to a_s, it is no longer possible for the agents to remain indifferent to interference. They must find cooperatively (k, f) pairs which lead to social optimality. Past a_s, simple Luenberger type observers and constant state estimate feedback cannot possibly maintain stability.

Let $a_m(f)$ be the maximum value of a such that $(K^*(a), f) \in S(a)$ and $a_m = \sup_f a_m(f)$. Then up to a_m, given the optimal isolated Kalman gain $K^*(a)$ and a pair $(K^*(a), f)$ within the stability region, it is possible to reverse engineer the choice of parameter $r > 0$ in the individual agent cost function (1.7) so that $[K^*(a), f]$ be an optimum isolated agent vector of control and estimation gains for some r_{Rev} iff

$$0 < r_{Rev} = \frac{-b(a-bf)}{f[a(a-bf)-1]} < \infty \tag{1.15}$$

with $a \neq 0$. The main results are summarized in Table 1.1.

Table 1.1 The boundaries of stability regions of separated gains

a	r	(K, f)
$0 < a < a_{Nash}$	$0 < r$	$(K^*(a), f^*(a, r)) \in S(a)$
$a_{Nash} < a < a_{\sup}$	$0 < r_{\inf}(a) < r$	$(K^*(a), f^*(a, r)) \in S(a)$
$a_{\sup} < a < a_m$	$0 < r_{Rev} = \frac{-b(a-bf)}{f[a(a-bf)-1]} < \infty$	$(K^*(a), f) \in S(a)$
$a_m < a < a_s$	–	$(K, f) \in S(a)$
$a_s < a$	–	$(K, f) \notin S(a)$

1.4.3 Deriving the exact optimal decentralized filter under the class of certainty equivalent constant feedback controllers

Since naive filtering reaches its stabilization capability (i.e. keeping costs finite), we explore a class of controllers which includes the class utilized in the previous part as a special case: that of certainty equivalent controllers under constant state feedback where the state estimate is computed exactly. It turns out no sufficient statistic is available, and all information must be kept to produce the best estimate. This means growing dimension filters. We compute the expressions of the latter and carry out the calculations in a semi recursive manner. We also explore finite memory implementation of the filters and the corresponding expressions.

Important observations:

— The proposed estimator (hereon also referred to as bulk filter) in combination with an arbitrary (stabilizing under perfect state observations) state estimate feedback gain, succeeds in maintaining the boundedness of the closed loop system even when individual systems are highly unstable.

— Existence of a stationarity threshold $\bar{a}(f, N)$ past which, the optimal filter gains never stationarize, i.e. remain time-varying, and essentially periodic in the case of weakly unstable agents.

— The steady state bulk filter with mostly a few coefficients non zero could be recovered by applying the naive Kalman filter equivalent equations.

— Persistently time varying behavior for large enough a

— Non-optimality of the Riccati gain (in general)

— The stabilizing capability of finite memory filtering approximations where only the last n measurements are preserved, improves by increasing the memory length n.

1.4.4 Computation of Nash equilibria for special classes of output feedback control laws

The complexity of certainty equivalent controllers precludes for the time being analysis of their stabilization abilities. However, inspired by the behavior of the filter, we explore necessary conditions (fixed point equations) that would lead to Nash equilibria under a class of total measurements preserving time varying output feedback stabilizing controllers, as N goes to infinity, for a finite length 2 time horizon. Subsequently the corresponding expressions are written for horizon larger than 2, by preserving only the contribution of the most recent 2 measurements. This leads to a heuristic estimate of Nash equilibria under a class of output time varying controllers with only recollection of the two most recent measurements.

1.5 Structure of The Thesis

The rest of this dissertation is organized as follows. Chapter 2 presents a critical review of the pertinent literature in the context of the defined problems. Chapter 3 presents the process for the research project as a whole and the general organization of the document, indicating the coherence of the articles in relation to the research goals. Chapters 4, 5 and 6, present the main results of this dissertation, which respectively, include Article 1, Article 2 and Article 3. Chapter 7 discusses the methodological aspects and results of this dissertation linked with the critical literature review. Concluding remarks and suggestions for future developments are stated in Chapter 8.

CHAPTER 2 CRITICAL LITERATURE REVIEW

The literature on the subject of this thesis has expanded in a number of directions. In this chapter, we provide a survey of those research directions.

2.1 Distributed estimation and control of multi-agent systems

We note that there is a vast literature on methods for distributed estimation and control of multi-agent systems. Here we refer to some of them. In particular, a framework for the design of collective behaviors for groups of identical mobile agents is described in Yang *et al.* (2008). Their approach is based on decentralized simultaneous estimation and control, where each agent communicates with neighbors and estimates the global performance properties of the swarm needed to make a local control decision. The survey paper Garin and Schenato (2010) refers to some works on distributed estimation and control applications using linear consensus algorithms up to 2010. The authors in Le Ny *et al.* (2011) have considered an attention-control problem in continuous time, which consists of scheduling sensor/target assignments and running the corresponding Kalman filters. In Olfati-Saber and Jalalkamali (2012), a theoretical framework for coupled distributed estimation and motion control of mobile sensor networks is introduced for collaborative target tracking (see also the references therein). In Roshany-Yamchi *et al.* (2013) a novel distributed Kalman filter algorithm along with a distributed model predictive control (MPC) scheme for large-scale multi-rate systems is proposed, where the decomposed multi-rate system consists of smaller subsystems with linear dynamics that are coupled via states. In this scheme, multiple control and estimation agents each determine actions for their own parts of the system. Via communication and in a cooperative way, the agents can take one another's actions into account. Moreover, estimation of population systems has been addressed in Ruess *et al.* (2011), where the authors proposed a novel method for estimating the moments of chemically reacting systems. Their method is based on closing the moment dynamics by replacing the moments of order $n + 1$ by estimates calculated from a small number of stochastic simulation runs. The resulting stochastic system is then used in an extended Kalman filter, where estimates of the moments of order up to n, obtained from the same simulation, serve as outputs of the system. Mahajan and Nayyar (2015) investigated decentralized LQG systems with partial history sharing information structure and identified finite-dimensional sufficient statistics for such systems.

2.2 Large population game theoretic models and mean field games

The very early interest in large population game theoretic models was already present in the book by Neumann and Morgenstern (1953). Also, a general equilibrium featuring a continuum of agents was presented in Aumann (1964). Since then there has been a vast literature on such models (see Khan and Sun (2002), Carmona (2004) and the references therein). Large-scale stochastic games with unbounded costs were studied in Adlakha *et al.* (2008). Since 2003, the mean field games (MFG) (or Nash certainty equivalence (NCE)) theory has been developed as a decentralized methodology for large population stochastic dynamic games with mean field couplings in their individual dynamics and cost functions, by Huang together with Caines and Malhamé (see Huang *et al.* (2007); Huang *et al.* (2003); Huang *et al.* (2006b), among others), and independently, in the context of partial differential equations and viscosity solutions by Lions and Lasry (see Lasry and Lions (2007); Lasry and Lions (2006a); Lasry and Lions (2006b)). In Li and Zhang (2008), the mean field linear quadratic Gaussian (LQG) framework was extended to systems of agents with Long Time Average (LTA) (i.e., ergodic) cost functions such that the set of control laws possesses an almost sure (a.s.) asymptotic Nash equilibrium property. The survey papers Buckdahn *et al.* (2011); Gomes and Saúde (2014); Caines (2015) provide an overview on the mean field game theory and also refer to some related works in this field. Moreover, the MFG methodology has been applied to wireless power control in Huang *et al.* (2003); Tembine *et al.* (2010); Aziz and Caines (2017), to coupled nonlinear oscillators subject to random disturbances in Yin *et al.* (2012), to control of a large number of electric water heating loads in Kizilkale and Malhamé (2014, 2016), to large population of Plug-in Electric Vehicles in Ma *et al.* (2013); Zhu *et al.* (2016), and to some models in economics such as in Weintraub *et al.* (2008); Gomes *et al.* (2016).

2.3 Stochastic non cooperative games with partial observation

The state estimation problem has been a fundamental and challenging problem in theory and applications of control systems. Stochastic non cooperative games with partial observation have been of interest since the late 1960s. LQG continuous-time zero-sum stochastic games with output measurements corrupted by additive independent white Gaussian noise were studied in Rhodes and Luenberger (1969a,b) under the constraint that each player is limited to a linear state estimator for generating its optimal controls. These results were extended to nonzero-sum non cooperative games in Saksena and Cruz (1982). In these works the authors assumed that the separation principle holds. In Kian *et al.* (2002), discrete-time nonzero-sum LQG non cooperative games with constrained state estimators and two different information structures were investigated, where it is shown that the optimal control laws do not satisfy the separation principle and the estimator characteristics depend on the controller gains.

Distributed decision-making with partial observation for large population stochastic multi-agent systems was studied in Caines and Kizilkale (2013, 2014, 2016); Firoozi and Caines (2016); Huang *et al.* (2006a); Şen and Caines (2016); Wang and Zhang (2013), where the synthesis of Nash strategies is investigated for the agents that are weakly coupled through either individual dynamics or costs. A new formulation of particle filters inspired by the mean-field game theoretic framework of Huang *et al.* (2007) was proposed by Yang *et al.* (2011b), Yang *et al.* (2011a) where each particle is equipped with a control to minimize a coupling cost. The authors in Pequito *et al.* (2011) introduced Mean Field Games (MFG) as a framework to develop an estimator for a class of nonlinear systems, where the goal is to minimize the expected minimum energy constrained to a consensus among all the possible evolutions given the initial conditions.

2.4 Networked multi-agent games

Game theory has emerged as a well-established discipline capable of providing a resourceful and effective framework for addressing control of large scale and distributed networked systems. In Eksin *et al.* (2014), a repeated network game where agents have quadratic utilities that depend on information externalities (an unknown underlying state) as well as payoff externalities (the actions of all other agents in the network) is considered. Agents play Bayesian Nash Equilibrium strategies with respect to their beliefs on the state of the world and the actions of all other nodes in the network. These beliefs are refined over subsequent stages based on the observed actions of neighboring peers. The authors introduce the Quadratic Network Game (QNG) filter that agents can run locally to update their beliefs, select corresponding optimal actions, and eventually learn a sufficient statistic of the network's state. They demonstrate the application of QNG filter on a Cournot market competition game and a coordination game to implement navigation of an autonomous team. Moon and Başar (2014) considered discrete-time linear quadratic-Gaussian (LQG) mean field games over unreliable communication links, where the individual dynamical system for each agent is subject to packet dropping. Salehisadaghiani and Pavel (2016) present an asynchronous gossip-based algorithm for finding a Nash equilibrium (NE) of a game in a distributed multiplayer network in such a way that players make decisions based on estimates of the other players' actions obtained from local neighbors. In Lian *et al.* (2017), to enable economically viable communication for multi-agent networked linear dynamic systems, a game-theoretic framework is proposed under the communication cost, or sparsity constraint, given by the number of communicating state/control input pairs. As this constraint tightens, the system transitions from dense to sparse communication, providing the trade-off between dynamic system performance and information exchange. Moreover, using the proposed sparsity constrained distributed social optimization and noncooperative game algorithms, they developed a method to allocate the costs of the communication infrastructure fairly and according to the agents' diverse needs for feedback and cooperation.

CHAPTER 3 PROCESS FOR THE RESEARCH PROJECT AS A WHOLE AND GENERAL ORGANIZATION OF THE DOCUMENT INDICATING THE COHERENCE OF THE ARTICLES IN RELATION TO THE RESEARCH GOALS

The research project of this thesis was initialized by an effort to generalize the mean field game theory (see Huang *et al.* (2007), for example) by possible dualization of mean field control results from control to estimation situations. In particular, in Abedinpour Fallah *et al.* (2013b) (see Appendix A), we proposed a distributed multi-agent decision-making with partial observation by employing the mean-field theory based on the work by Huang *et al.* (2006a). Under the assumption of rationality for each agent (where an agent is rational if it optimizes its own cost while assuming rationality of other agents in optimizing their costs) and using the state aggregation technique, our proposed Algorithm combines the Kalman filtering for state estimation and the linear quadratic Gaussian (LQG) feedback controller. For large N, as in the typical mean field analysis, we assumed in the first place that conditions are satisfied so that controlled agents become asymptotically independent (in the limit of large population), and furthermore the coupling term (mass effect) is approximated by a deterministic continuous function (to be determined later). This leads to uncoupled measurement equations, and therefore the optimal state estimation would be given by the standard scalar Kalman filtering. Therefore, in Article 1 presented in Chapter 4, for a finite number of agents, we establish that provided some constraints on cost and dynamic parameters are satisfied, optimal separated control laws obtained by ignoring the interference coupling, are asymptotically as the number of agents goes to infinity, individually optimal. Moreover, they induce a Nash equilibrium, while and for finite N, they induce an ϵ-Nash equilibrium. More generally though, optimal separated control laws may not be asymptotically optimal, and can in fact result in unstable overall behavior. Thus we consider a class of parameterized decentralized control laws whereby the separated Kalman gain is treated as the arbitrary gain of a Luenberger like observer. The nature of optimal cooperative control policies within the considered class is explored. Note also that even within the cooperative social setting, we establish an upper limit on the degree of instability of individual agent systems, before stabilization via feedback gains on Luenberger like state estimates becomes impossible.

Since naive filtering reaches its stabilization capability (i.e. keeping costs finite), in Article 2 presented in Chapters 5, we explore a class of controllers which includes the class utilized in the previous Chapter as a special case: that of certainty equivalent controllers under constant state feedback where the state estimate is computed exactly. It turns out no sufficient statistic is available, and all information must be kept to produce the best estimate. This means growing dimension filters. We compute the expressions of the latter and carry out the calculations in a semi recursive manner. We also explore finite memory implementation of the filters and the corresponding expressions.

The complexity of certainty equivalent controllers precludes for the time being analysis of their stabilization abilities. However, inspired by the behavior of the filter, in Article 3 presented in Chapter 6, we explore necessary conditions (fixed point equations) that would lead to Nash equilibria under a class of total measurements preserving time varying output feedback stabilizing controllers, as N goes to infinity, for a finite length 2 time horizon. Subsequently the corresponding expressions are written for horizon larger than 2, by preserving only the contribution of the most recent 2 measurements. This leads to a heuristic estimate of Nash equilibria under a class of output time varying controllers with only recollection of the two most recent measurements.

Table 3.1 provides an overview of these studies, their methodological approaches and their main results.

Table 3.1 An overview of the articles

Article	Approaches and their main results
1	Constant state feedback and Luenberger like naive observer with $h = 0$
1	Exploring limits of naive Kalman/Luenberger based Nash equilibria
1	Isolated (naive) policies can stabilize up to $a < a_s$
2	Stabilization via exact and approximate growing dimension filters
2	Boundedness of the closed loop system with highly unstable individual systems
2	Persistently time varying behavior for large enough a
2	Non-optimality of the Riccati gain (in general)
3	A heuristic estimate of Nash eq. under a classe of output feedback control laws
3	With only recollection of the two most recent measurements
3	A fixed point based algorithm
3	A forward sweep for state estimation
3	A backwards sweep for dynamic programming based optimization
4	An article on the continuous-time version of the models
4	Using the mean field game theory and state aggregation technique

CHAPTER 4 ARTICLE 1 : A CLASS OF INTERFERENCE INDUCED GAMES : ASYMPTOTIC NASH EQUILIBRIA AND PARAMETERIZED COOPERATIVE SOLUTIONS

Mehdi Abedinpour Fallah, Roland P. Malhamé and Francesco Martinelli
Automatica, vol. 69, pp. 181-194, 2016.

4.1 Abstract

We consider a multi-agent system with linear stochastic individual dynamics, and individual linear quadratic ergodic cost functions. The agents partially observe their own states. Their cost functions and initial statistics are a priori independent but they are coupled through an interference term (the mean of all agent states), entering each of their individual measurement equations. While in general for a finite number of agents, the resulting optimal control law may be a non linear function of the available observations, we establish that for certain classes of cost and dynamic parameters, optimal separated control laws obtained by ignoring the interference coupling, are asymptotically optimal when the number of agents goes to infinity, thus forming for finite N, an ϵ-Nash equilibrium. More generally though, optimal separated control laws may not be asymptotically optimal, and can in fact result in unstable overall behavior. Thus we consider a class of parameterized decentralized control laws whereby the separated Kalman gain is treated as the arbitrary gain of a Luenberger like observer. System stability regions are characterized and the nature of optimal cooperative control policies within the considered class is explored. Numerical results and an application example for wireless communications are reported.

4.2 Introduction

There has been a surge of interest in the study and analysis of large population stochastic multi-agent systems due to their wide variety of applications over the past several years. Many practical applications and examples of these systems arise in engineering, biological, social and economic fields, such as wireless sensor networks (Chong and Kumar (2003)), very large scale robotics (Reif and Wang (1999)), controlled charging of a large population of electric vehicles (Karfopoulos and Hatziargyriou (2013)), synchronization of coupled oscillators (Yin *et al.* (2012)), swarm and flocking phenomenon in biological systems (Grönbaum and Okubo (1994); Passino (2002)), evacuation of large crowds in emergency situations (Helbing *et al.* (2000); Lachapelle (2010)), sharing and competing for resources on the Internet (Altman *et al.* (2006)), to cite a few. Large-scale stochastic games with unbounded costs were studied in Adlakha *et al.* (2008). Mean field game theory, which addresses a class of dynamic games with a large number of agents in which each agent interacts with the average or so-called mean field effect of other agents via couplings in their individual dynamics and cost functions, was studied in Huang *et al.* (2007, 2012); Lasry and Lions (2007); Nourian *et al.* (2012); Wang and Zhang (2012, 2014). In Li and Zhang (2008), the mean field linear quadratic Gaussian (LQG) framework was extended to systems of agents with Long Time Average (LTA) (i.e., ergodic) cost functions such that the set of control laws possesses an almost sure (a.s.) asymptotic Nash equilibrium property.

Stochastic Nash games with partial observation have been of interest since the late 1960s. LQG continuous-time zero-sum stochastic games with output measurements corrupted by additive independent white Gaussian noise were studied in Rhodes and Luenberger (1969a,b) under the constraint that each player is limited to a linear state estimator for generating its optimal controls. These results were extended to nonzero-sum Nash games in Saksena and Cruz (1982). In these works the authors assumed that the separation principle holds. In Kian *et al.* (2002), discrete-time nonzero-sum LQG Nash games with constrained state estimators and two different information structures were investigated, where it is shown that the optimal control laws do not satisfy the separation principle and the estimator characteristics depend on the controller gains.

Distributed decision-making with partial observation for large population stochastic multi-agent systems was studied in Caines and Kizilkale (2013, 2014); Huang *et al.* (2006a); Wang and Zhang (2013), where the synthesis of Nash strategies is investigated for the agents that are weakly coupled through either individual dynamics or costs. In Abedinpour Fallah *et al.* (2013a,b, 2014) the authors studied a somewhat dual situation whereby large populations of partially observed stochastic agents, although a priori individually independent, are coupled only via their observation structure. The latter involves an interference term depending on the empirical mean of all agent states. The study of such measurement-coupled systems is inspired by a variety of applications, including for instance the communications model for power control in cellular telephone systems (Huang *et al.* (2004); Perreau and Anderson (2006)), where any conversation in a cell acts as interference on the other conversations in that cell. Indeed, despite the so-called signal processing gain achieved thanks to a user's specific coding advantage (and considered in our model to be of order $1/N$ where N is the total number of agents), the ability of the base station to correctly decode the signals sent by a given mobile, remains limited by interference formed by the superposition of all other in cell user signals. Viewed in this light, the studied problem can be considered as a game over a noisy channel.

Individual agent dynamics are assumed to be linear, stochastic, with linear local state measurements, and in the current paper, we focus on the case where the measurements interaction model is assumed to depend only on the empirical mean of agents states in a purely additive manner. In general, in such decentralized control problems, the measurement system could be used for some sort of signalling, and control and estimation are typically coupled (Witsenhausen (1968)). We assume that each agent is constrained to use a linear Kalman filter-like state estimator to generate its optimal strategies. For a finite number of agents, we establish that for certain classes of cost and dynamic parameters, optimal separated control laws obtained by ignoring the interference coupling, are asymptotically optimal when the number of agents goes to infinity, thus forming for finite N, an ϵ-Nash equilibrium. More generally though, optimal separated control laws may not be asymptotically optimal, and can in fact result in unstable overall behavior. Thus we consider a class of parameterized decentralized control laws whereby the separated Kalman gain is treated as the arbitrary gain of a Luenberger like observer. System stability regions are characterized and the nature of optimal cooperative control policies within the considered class is explored.

The rest of the paper is organized as follows. The problem is defined and formulated in Section 4.3. Section 4.4 presents the closed-loop dynamics model. In Section 4.5, a decentralized control and state estimation algorithm via stability analysis is described and a characterization of its optimality properties is given. Section 4.6 presents parameterized cooperative solutions. Also, both Section 4.5 and Section 4.6 provide some numerical simulation results. Section 4.7 presents an application example for wireless communications. Concluding remarks are stated in Section 4.8.

4.3 Problem formulation

Consider a system of N agents, with individual scalar dynamics for simplicity of computations. The evolution of the state component is described by

$$x_{k+1,i} = ax_{k,i} + bu_{k,i} + w_{k,i} \tag{4.1}$$

with partial scalar state observations given by:

$$y_{k,i} = cx_{k,i} + h\left(\frac{1}{N}\sum_{j=1}^{N} x_{k,j}\right) + v_{k,i} \tag{4.2}$$

for $k \geq 0$ and $1 \leq i \leq N$, where $x_{k,i}, u_{k,i}, y_{k,i} \in \mathbb{R}$ are the state, the control input and the measured output of the i^{th} agent, respectively. The random variables $w_{k,i} \sim \mathcal{N}(0, \sigma_w^2)$ and $v_{k,i} \sim \mathcal{N}(0, \sigma_v^2)$ represent independent Gaussian white noises at different times k and at different agents i. The Gaussian initial conditions $x_{0,i} \sim \mathcal{N}(\bar{x}_0, \sigma_0^2)$ are mutually independent and are also independent of $\{w_{k,i}, v_{k,i}, 1 \leq i \leq N,\ k \geq 0\}$. σ_w^2, σ_v^2 and σ_0^2 denote the variance of $w_{k,i}$, $v_{k,i}$ and $x_{0,i}$, respectively. Moreover, a is a scalar parameter and $b, c, h > 0$ are positive scalar parameters.

The problem to be considered is to synthesize the linear time invariant decentralized separated policies such that each agent is stabilized by a feedback control of the form

$$u_{k,i} = -f\hat{x}_{k,i}, \tag{4.3}$$

where $\hat{x}_{k,i}$ is an estimator of $x_{k,i}$ based only on local observations of the i^{th} agent, and f is a constant scalar gain. For the purposes of the paper, the class of *decentralized separated policies* (4.3) includes all control policies satisfying the following three conditions : (i) they are defined by two time invariant feedback gains K and f, (ii) they are separated in that the control is a linear feedback $-f\hat{x}_{k,i}$ on the state estimate of $x_{k,i}$, while the state estimate $\hat{x}_{k,i}$ is obtained from a Luenberger like observer equation under the assumed state estimate feedback structure, i.e., it evolves according to:

$$\hat{x}_{k+1,i} = (a - bf)\hat{x}_{k,i} + K(y_{k+1,i} - c(a - bf)\hat{x}_{k,i}), \tag{4.4}$$

(iii) they are decentralized in that the state estimate is based solely on agent based observations $y_{k,i}$.

Furthermore, when the gain K is the Kalman gain as obtained when assuming zero interference in the local measurements (setting $h = 0$ in (4.2)), the resulting estimator (4.4) will be called the *naive* Kalman filter. Moreover, the individual cost function for each agent is given by

$$J_i \triangleq \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2). \tag{4.5}$$

where $\overline{\lim}$ denotes the limit superior and $r > 0$ is a positive scalar parameter.

Assumption 1. *To simplify the synthesis procedure we assume zero mean for initial condi-*

tions of all agents, i.e., $\mathbb{E}x_{0,i} = \bar{x}_0 = 0, i \geq 1$.

Remark 4.1. *To show that the decentralized control problem formulated here is a game, let us assume for the sake of discussion that the original agent dynamics is unstable. Then it suffices to observe that, for finite N at least, the inability of a single agent to stabilize its own dynamics would have direct consequences on the ability of other agents to stabilize their own, hence demonstrating the impact of that agent on other agents' individual costs.*

4.4 Closed-loop dynamics model

4.4.1 Closed-loop agent dynamics

In this section first we obtain the 4^{th} order model of the closed-loop agent dynamics. In particular, when local state estimate feedback (4.3) is included in the i^{th} agent state equation (4.1), the result is as follows:

$$x_{k+1,i} = ax_{k,i} - bf\hat{x}_{k,i} + w_{k,i}. \tag{4.6}$$

In addition, anticipating the need to account for the influence of average states in the dynamics through the measurement equation, and letting a tilde $(\tilde{.})$ indicate an averaging over agents operation, we define:

$$m_k = \frac{1}{N}\sum_{j=1}^{N} x_{k,j}, \quad \tilde{m}_k = \frac{1}{N}\sum_{j=1}^{N} \hat{x}_{k,j}, \tag{4.7}$$

$$\tilde{w}_k = \frac{1}{N}\sum_{j=1}^{N} w_{k,j}, \quad \tilde{v}_k = \frac{1}{N}\sum_{j=1}^{N} v_{k,j}. \tag{4.8}$$

Now, combining (4.6), (4.7), (4.8), we obtain the population average state evolution:

$$m_{k+1} = am_k - bf\tilde{m}_k + \tilde{w}_k. \tag{4.9}$$

Also averaging the estimate $\hat{x}_{k+1,i}$ given by (4.4), yields the population average state estimate dynamics:

$$\tilde{m}_{k+1} = (a - bf)\tilde{m}_k + K((c + h)m_{k+1} - c(a - bf)\tilde{m}_k + \tilde{v}_{k+1}). \tag{4.10}$$

Thus, combining (4.4) and (4.6) with (4.9) and (4.10) yields

$$X_{k+1,i} = AX_{k,i} + DW_{k,i}, \tag{4.11}$$

where the augmented state is

$$X_{k,i} = [x_{k,i}, \hat{x}_{k,i}, m_k, \tilde{m}_k]^T, \tag{4.12}$$

and matrix A is given by

$$A = \begin{bmatrix} a & -bf & 0 & 0 \\ acK & a(1-cK)-bf & ahK & -bfhK \\ 0 & 0 & a & -bf \\ 0 & 0 & a(c+h)K & a_{4,4} \end{bmatrix}, \tag{4.13}$$

with

$$a_{4,4} = a(1-cK) - bf(1+hK), \tag{4.14}$$

and also we have:

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ cK & hK & K & 0 \\ 0 & 1 & 0 & 0 \\ 0 & (c+h)K & 0 & K \end{bmatrix}, \quad W_{k,i} = \begin{bmatrix} w_{k,i} \\ \tilde{w}_k \\ v_{k+1,i} \\ \tilde{v}_{k+1} \end{bmatrix}. \tag{4.15}$$

Furthermore, the covariance matrix of $W_{k,i}$ is given by:

$$\Sigma_w = \begin{bmatrix} \sigma_w^2 & \frac{\sigma_w^2}{N} & 0 & 0 \\ \frac{\sigma_w^2}{N} & \frac{\sigma_w^2}{N} & 0 & 0 \\ 0 & 0 & \sigma_v^2 & \frac{\sigma_v^2}{N} \\ 0 & 0 & \frac{\sigma_v^2}{N} & \frac{\sigma_v^2}{N} \end{bmatrix}. \tag{4.16}$$

4.4.2 Population average dynamics

The mean state and mean state estimate equation can be isolated from (4.11) as:

$$\begin{bmatrix} m_{k+1} \\ \tilde{m}_{k+1} \end{bmatrix} = A_p \begin{bmatrix} m_k \\ \tilde{m}_k \end{bmatrix} + D_p \begin{bmatrix} \tilde{w}_k \\ \tilde{v}_{k+1} \end{bmatrix}, \tag{4.17}$$

where

$$A_p = \begin{bmatrix} a & -bf \\ a(c+h)K & a(1-cK)-bf(1+hK) \end{bmatrix}, \; D_p = \begin{bmatrix} 1 & 0 \\ (c+h)K & K \end{bmatrix}. \tag{4.18}$$

Also, the covariance matrix of $[\tilde{w}_k, \tilde{v}_{k+1}]^T$ is given by:

$$\Sigma_{w,p} = \begin{bmatrix} \frac{\sigma_w^2}{N} & 0 \\ 0 & \frac{\sigma_v^2}{N} \end{bmatrix}. \tag{4.19}$$

4.5 Decentralized controller and state estimator

4.5.1 The race between N and T

It may appear obvious that as N goes to infinity, from Assumption 1 and (4.7) we have $\mathbb{E}[m_k] = 0$, and as a result at least asymptotically, the agent systems become essentially independent and individually *optimal control laws* are obtained via a Kalman filter K^* coupled with a gain f^* obtained from a Riccati equation. However, it turns out that while this is indeed correct if N is allowed to go to infinity before the length of the control horizon T is, *it is no longer always true if instead T is allowed to go to infinity first.* Theorem 4.1 establishes the separation result when N goes to infinity before T. Thus, in general, interchanging the orders of limits in N and T does not produce the same results.

Theorem 4.1. *In the case of the infinite population limit, the separated optimal policies consisting of the naive Kalman filter (4.4) denoted as $K^*(a)$, and the control gain obtained by the appropriate algebraic Riccati equation denoted as $f^*(a, r)$, define the optimal solution.*

Proof: In the case of the infinite population limit, the agent i observes that given independence and $\mathbb{E}[x_{0,j}] = 0$ for $j \geq 1$, $\frac{1}{N}\sum_{j=1}^{N} x_{1,j} \sim 0$ a.e. ($\sim$ a.e. means converges pointwise almost everywhere). The agent makes the assumption that this situation will persist in the future and under this most optimistic assumption computes its optimal K^* and f^* based control law. At step 2, because the applied control inputs are independent from one agent to the other, the agents states remain independent, and since the optimal control law is stabilizing, the individual state variance remains bounded while the mean is still zero. Once again then $\frac{1}{N}\sum_{j=1}^{N} x_{2,j} \sim 0$, and the initial i^{th} agent optimistic assumption is validated. In general, by assuming that the separated optimal control law is applied and that at step k the agents have zero mean independent states, one can establish that the property of zero mean state independence persists at step $(k+1)$, thus yielding $\frac{1}{N}\sum_{j=1}^{N} x_{(k+1),j} \sim 0$. As a result, one can recursively establish that $\forall k$, under the K^* and f^* based control law, the optimistic assumption $\frac{1}{N}\sum_{j=1}^{N} x_{k,j} \sim 0$ holds, while under that assumption K^* and f^* would indeed be parameters of the optimal control law. Since no improvement to estimation can occur for N infinite, if $\frac{1}{N}\sum_{j=1}^{N} x_{k,j} \neq 0$, then K^* and f^* will indeed define the optimal control law.$\square$

While one can suspect that K^* and f^* will be asymptotically optimal in a number of situations, it is not always so because as it turns out, stability is at the heart of the question. In order to illustrate this point, we show a simulation of the behavior of an agent sample path and the agents mean sample path for increasing values of N when $a = 2.5$, $r = 1$, and the optimal separated gains $K^*(a)$, $f^*(a, r)$ are applied. As established later in Section 4.2.1, for $a = 2.5$, and $r = 1$, the couple $(K^*(a), f^*(a, r))$ is outside the mean square stability region of the mean state, thus explaining the unstable behavior of the agents states observed in Figures 4.1 and 4.2. As observed, in this case a larger but *finite* N, can only *delay* but cannot indefinitely push away the onset of instability.

Numerical results I

The numerical results reported in this paper are obtained considering the following parameter setting: $b = 1$, $c = 1$, $h = 1$, $\sigma_v = 1$, $\sigma_w = 1$ and initial standard deviation $\sigma_0 = 1$. The value of a and f (or $a_f = a - bf$) will be specified in the different simulations. Figs. 4.1–4.2 which are obtained using the optimal Kalman-Riccati pair, show a simulation case where the cost runs to infinity for $N = 1000$ and $N = 1000000$, respectively.

4.5.2 Stability analysis

It is desirable to investigate the necessary and sufficient conditions for the closed-loop individual systems to be stable. By applying the Jury's stability criterion Ogata (1995) to a second order polynomial we firstly note that the following lemma holds.

Lemma 4.1. *(Ritzerfeld (2005)) A second-order discrete-time linear system having the following characteristic polynomial*

$$a_0 z^2 + a_1 z + a_2, \tag{4.20}$$

with real coefficients a_0, a_1 and a_2, is stable iff

$$a_0 + a_1 + a_2 > 0, \tag{4.21}$$

$$a_0 - a_1 + a_2 > 0, \tag{4.22}$$

$$a_0 - a_2 > 0. \tag{4.23}$$

Also, we have the following Lemma.

Lemma 4.2. *The population average dynamics (4.17) is such that the pair $(A_p, D_p\sqrt{\Sigma_{w,p}})$ is controllable.*

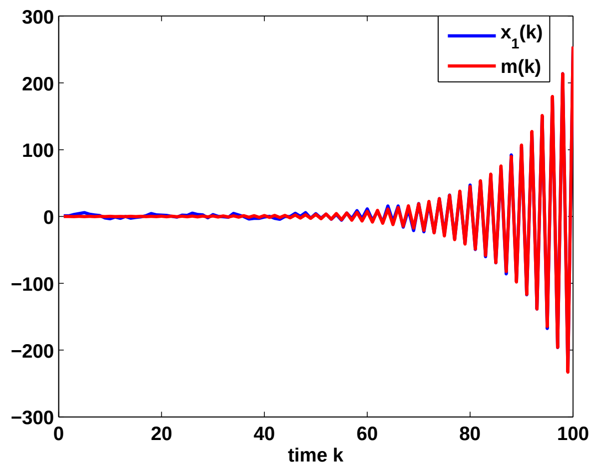

Figure 4.1 Unstable behavior of agent 1 and of the average of all the population when $a = 2.5$, $r = 1$, and $N = 1000$.

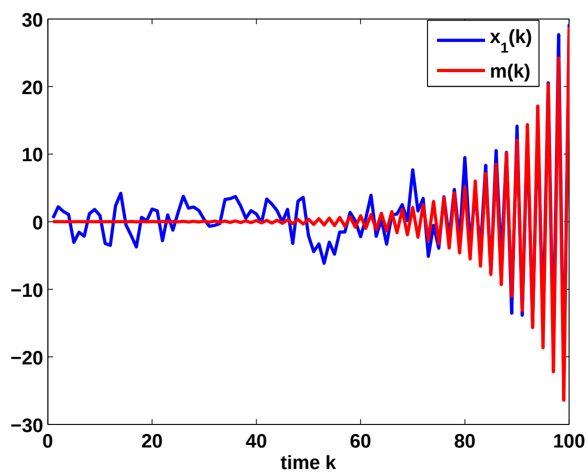

Figure 4.2 Unstable behavior of agent 1 and of the average of all the population when $a = 2.5$, $r = 1$, and $N = 1000000$.

Proof: It is not difficult to show that $D_p\sqrt{\Sigma_{w,p}}$ is invertible. Therefore, the controllability matrix associated with the pair $(A_p, D_p\sqrt{\Sigma_{w,p}})$ is full rank. □

Now the mean system mean square stability conditions are given by the following theorem (see Xie and Khargonekar (2012) and the references therein for definition of mean square stability).

Theorem 4.2. *Mean system (4.17) is mean square stable (MSS) iff the following inequalities are satisfied:*

$$K(ac(1-a)+bf(ac+h)) > bf(a-1)-(1-a)^2, \tag{4.24}$$

$$K(-ac(1+a)+bf(ac-h)) > bf(a+1)-(1+a)^2, \tag{4.25}$$

$$|a(a-bf)(1-cK)| < 1. \tag{4.26}$$

Proof: First note that since by Lemma 4.2, system (4.17) is controllable by noise, a necessary and sufficient condition for the mean square stability of $[m_k, \tilde{m}_k]^T$ is the stability of matrix A_p (see Theorem 3.13, p. 31 in Kumar and Varaiya (1986)). Now A_p will be stable iff the Jury stability test (Ogata (1995)) is met. In particular, the characteristic polynomial of (4.17) is given by:

$$z^2 + (bf(1+hK) - a(2-cK))z + a(a-bf)(1-cK). \tag{4.27}$$

Thus, by applying Lemma 4.1 with

$$a_0 = 1, \tag{4.28}$$

$$a_1 = bf(1+hK) - a(2-cK), \tag{4.29}$$

$$a_2 = a(a-bf)(1-cK), \tag{4.30}$$

the theorem is proved. □

The next theorem gives individual state mean square stability conditions.

Theorem 4.3. *Individual agent systems described by (4.11) are MSS iff:*

$$K(ac(1-a)+bf(ac+h)) > bf(a-1)-(1-a)^2, \tag{4.31}$$

$$K(-ac(1+a)+bf(ac-h)) > bf(a+1)-(1+a)^2, \tag{4.32}$$

$$|a-bf| < 1, \tag{4.33}$$

$$|a(1-cK)| < 1. \tag{4.34}$$

Proof: We first note that, in view of the measurement structure of agent i, MSS of the mean population dynamics is necessary for the MSS of individual agents. Thus, inequalities (4.24)-(4.26) form necessary conditions. In addition since the pair $(A, D\sqrt{\Sigma_w})$ is controllable, by Theorem 3.13, p. 31 of Kumar and Varaiya (1986), the complete state (including individual agent state and agent state estimate) is MSS iff matrix A is stable. However, given the block triangular structure of matrix A, its eigenvalues are the union of those of the diagonal blocks. Therefore, in addition to inequalities (4.24)-(4.26), one must also satisfy (4.33)-(4.34) obtained from the Jury stability criterion for the upper block. The concatenation of all these inequalities leads to (4.31)-(4.34) as necessary and sufficient conditions for MSS of individual agent state and state estimate dynamics. $\square$

Remark 4.2. *For given a, the (K, f) stability region is independent of N because the stability conditions (4.24)-(4.26) and (4.31)-(4.34) are independent of N.*

The next lemma is about the stability region for f.

Lemma 4.3. *For given a, $K^*(a)$, the stability region for f, if non empty, is an interval $(f_{\inf}(a), f_{\sup}(a))$.*

Proof: This lemma is proved using the stability conditions of Theorem 4.3. In particular, under stability condition $|a(1-cK^*)| < 1$ and noting that $c, h > 0$, $K^* \geq 0$ we have:

$$-1+a(1-cK^*)-hK^* < 0, \tag{4.35}$$

$$1+a(1-cK^*)+hK^* > 0. \tag{4.36}$$

Moreover, condition (4.31) can be written as:

$$ac(1-a)K^* + (1-a)^2 > bf(-1+a(1-cK^*)-hK^*). \tag{4.37}$$

Thus, combining (4.35) and (4.37) yields

$$f > \frac{1-a(1-cK^*)}{1-a(1-cK^*)+hK^*}(\frac{a-1}{b}) \cdot \tag{4.38}$$

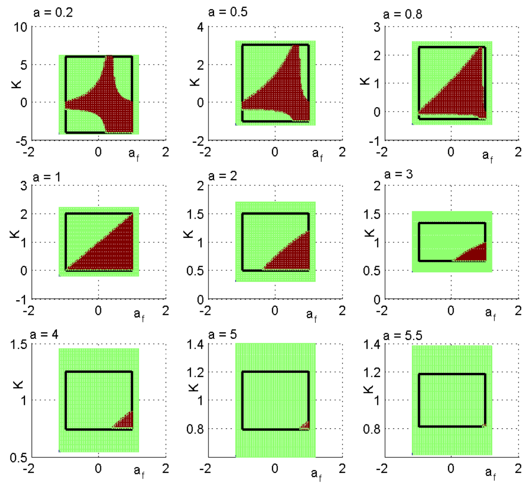

Figure 4.3 Stability regions: the box inside the black frame defines the region where (4.33) and (4.34) are met. This bounded area has been numerically explored to determine the stability regions (brown areas).

Similarly, from condition (4.32) and (4.36) we have:

$$f < \frac{1+a(1-cK^*)}{1+a(1-cK^*)+hK^*}\left(\frac{a+1}{b}\right) \cdot \tag{4.39}$$

Furthermore, condition (4.33) can be written as:

$$\frac{a-1}{b} < f < \frac{a+1}{b} \tag{4.40}$$

Note that (4.38), (4.39), and (4.40) have to be all satisfied simultaneously. Therefore, noting that

$$0 < \frac{1-a(1-cK^*)}{1-a(1-cK^*)+hK^*} < 1, \tag{4.41}$$

$$0 < \frac{1+a(1-cK^*)}{1+a(1-cK^*)+hK^*} < 1, \tag{4.42}$$

we have:

$$f_{\inf}(a) = \frac{(a-1)(1-a(1-cK^*))}{b(1-a(1-cK^*)+hK^*)}\,, \, if \quad a < 1; \tag{4.43}$$

$$f_{\inf}(a) = \frac{a-1}{b}\,, \quad if \quad a \geq 1; \tag{4.44}$$

$$f_{\sup}(a) = \frac{a+1}{b}\,, \quad if \quad a \leq -1; \tag{4.45}$$

$$f_{\sup}(a) = \frac{(a+1)(1+a(1-cK^*))}{b(1+a(1-cK^*)+hK^*)}\,, \, if \; a > -1. \tag{4.46}$$

□

Numerical results II

Fig. 4.3 is a representation of the stability regions defined by Theorem 4.3, and associated with the parameter set in Section 4.5.1 when a varies from $a = 0.2$ to $a = 5.5$. It is observed that the stability region gradually shrinks as a increases until it all but vanishes at $a = 5.5$.

4.5.3 Reverse engineering agent cost functions for stability

One of the main goals of the paper is to identify parameter sets for which asymptotically, as the number of agents increases without bound, separated optimal control policies (i.e., based on $K^*(a)$ and $f^*(a,r)$) become optimal for the measurements coupled agent system itself. Since stability is clearly a necessary condition for this to happen, to this end, we shall be concerned with the following question: Given parameter a, the optimal isolated Kalman gain $K^*(a)$, and a pair $(K^*(a), f)$ within the stability region, is it always possible to reverse engineer the choice of parameter $r > 0$ in the individual agent cost function (4.5) so that $f = f^*(a,r)$? In order to answer that question, we develop the following steps called Algorithm 1.

Algorithm 1

— Apply the naive Kalman filter

$$\hat{x}_{k+1,i} = (a-bf)\hat{x}_{k,i} + K(y_{k+1,i} - c(a-bf)\hat{x}_{k,i}),$$

with the steady-state scalar gain

$$K^*(a) = \frac{cP(a)}{c^2P(a)+\sigma_v^2}\,,$$

where P is the unique positive solution of

$$c^2P^2(a) + ((1-a^2)\sigma_v^2 - c^2\sigma_w^2)P(a) - \sigma_w^2\sigma_v^2 = 0.$$

— For some fixed a, let

$$u_{k,i} = -f\hat{x}_{k,i},$$

and find the possible values of stabilizing f using Theorem 4.3.

— For each one such stabilizing f, reverse engineer if possible the cost structure

$$J_i \triangleq \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2),$$

for parameter r by verifying if $0 < \frac{-b(a-bf)}{f[a(a-bf)-1]} < \infty$. If so, set

$$r = \frac{-b(a-bf)}{f[a(a-bf)-1]},$$

that is the unique value of r for which f is optimal.

Lemma 4.4. *For $a \neq 0$, $[K^*(a), f]$ is an optimum isolated agent vector of control and estimation gains for some r iff*

$$0 < r = \frac{-b(a-bf)}{f[a(a-bf)-1]} < \infty. \tag{4.47}$$

Proof: In order to reverse engineer the cost structure in Algorithm 1, we use the expression of the optimal feedback gain assumed to be f to express Σ,

$$\Sigma = \frac{rf}{b(a-bf)}, \tag{4.48}$$

the positive solution of the algebraic Riccati equation in

$$b^2\Sigma^2 + (r - a^2r - b^2)\Sigma - r = 0. \tag{4.49}$$

Then solving the resultant equation for the candidate r yields:

$$r = \frac{-b(a-bf)}{f[a(a-bf)-1]}. \tag{4.50}$$

Thus for $a \neq 0$, whenever the expression in the right-hand side of (4.50) is strictly positive and finite, the pair $(K^*(a), f)$ will be isolated agent optimal policies for the corresponding value of r. Note that for $a = 0$, the optimal feedback gain is zero, and any $r \in (0, \infty)$ will correspond to a potential reverse engineering cost solution. Except for the indetermination at $a = 0$, either the reverse cost solution does not exist, or it exists uniquely. □

Numerical results III

The reverse engineering regions are depicted in Fig. 4.4 and Fig. 4.5. It is observed that the stability limit on a is obtained as $|a| \leq 3.68$. Moreover, for all $a \in (-2.52, 2.52)$ and (a, a_f) in the stability region, where a_f is defined as $a_f = a - bf$, it is possible to reverse engineer a cost parameter r for isolated optimality. In particular, Fig. 4.5 confirms Lemma 4.3; in fact, for each a and $K^*(a)$, the stabilizing f belongs to an interval $(f_{inf}(a), f_{sup}(a))$. Lemma 4.4 has been used to evaluate if each stabilizing f corresponds to a positive value of r (i.e., if reverse engineering holds for that f).

4.5.4 Asymptotic optimality and ϵ-Nash equilibrium properties

In this section, we first establish conditions for the asymptotic optimality of isolated agent optimal gains.

Theorem 4.4. *Let $S(a)$ designate the (K, f) stability region associated with parameter a, and let $K^*(a)$ and $f^*(a, r)$, respectively, be the optimal estimation and control gains associated with the isolated agent optimal control problem (when $h = 0$, i.e., with zero average coupling term) for some $0 < r < \infty$. If $(K^*(a), f^*(a, r)) \in S(a)$, then the couple $(K^*(a), f^*(a, r))$ defines an asymptotically optimal policy for the coupled agents problem as $N \to \infty$. Furthermore, for N finite, it is ϵ-optimal with ϵ of order $1/N$ over any closed subset of $S(a)$ containing $(K^*(a), f^*(a, r))$, that is to say,*

$$J_i(K^*, f^*) - \epsilon \leq \inf_{(K,f) \in S(a)} J_i(K, f) \leq J_i(K^*, f^*). \tag{4.51}$$

Proof: We first note that for any $(K, f) \in S(a)$, from conditions (4.33) and (4.34), respectively, f and K will be bounded. Then using (4.3) and under the assumption that

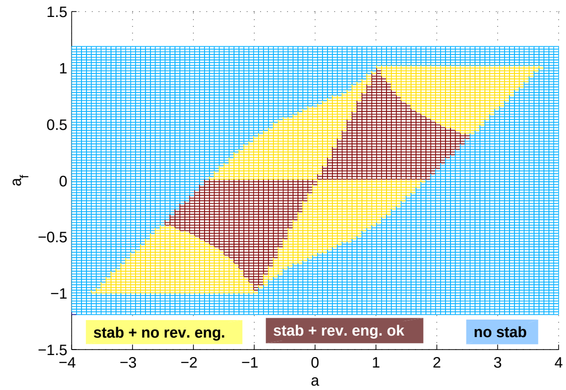

Figure 4.4 Stability and reverse engineering regions in (a, a_f)

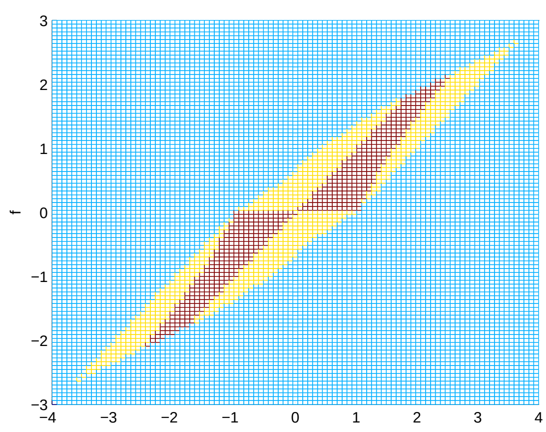

Figure 4.5 Stability and reverse engineering regions in (a, f) (same color meaning as Fig. 4.4)

$(K^*(a), f^*(a,r)) \in S(a)$, a direct calculation of the actual cost functions yields

$$
\begin{aligned}
J_i &= \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2), \\
&= \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(X_{k,i}^T\bar{Q}X_{k,i}), \\
&= \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}Tr(\bar{Q}X_{k,i}X_{k,i}^T), \\
&= Tr(\bar{Q}\bar{P}_\infty),
\end{aligned} \tag{4.52}
$$

where

$$
X_{k,i} = \begin{bmatrix} x_{k,i} \\ \hat{x}_{k,i} \\ m_k \\ \tilde{m}_k \end{bmatrix}, \bar{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & rf^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{4.53}
$$

$\bar{P}_k = \mathbb{E}[X_{k,i}X_{k,i}^T]$, and $\bar{P}_\infty = \lim_{k\to\infty}\bar{P}_k$. Note that $\bar{P}_\infty$ can be directly calculated from the covariance equation of the closed-loop system (4.11) and is given as the unique positive definite solution of the Lyapunov equation:

$$
A\bar{P}_\infty A^T - \bar{P}_\infty + D\Sigma_w D^T = 0, \tag{4.54}
$$

where Σ_w denotes the covariance matrice of $W_{k,i}$. Now Σ_w can be written as:

$$
\Sigma_w = \Sigma_{w_1} + \frac{1}{N}\Sigma_{w_2}, \tag{4.55}
$$

where

$$
\Sigma_{w_1} = \begin{bmatrix} \sigma_w^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_v^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Sigma_{w_2} = \begin{bmatrix} 0 & \sigma_w^2 & 0 & 0 \\ \sigma_w^2 & \sigma_w^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_v^2 \\ 0 & 0 & \sigma_v^2 & \sigma_v^2 \end{bmatrix}. \tag{4.56}
$$

Since (4.54) is a linear equation, superposition holds and $\bar{P}_\infty$ can be split into two components respectively, $\bar{P}_{\infty_1}$ and $\bar{P}_{\infty_2}$ corresponding to the respective contributions of Σ_{w_1} and $\frac{1}{N}\Sigma_{w_2}$. In addition, let the corresponding contributions to the total cost be respectively denoted

$\alpha(K, f)$ and $\frac{1}{N}\beta(K, f)$. Thus

$$J_i(K, f, \frac{1}{N}) = \alpha(K, f) + \frac{1}{N}\beta(K, f). \tag{4.57}$$

Since (K, f) belongs to $S(a)$ the nature of which is independent of N, then both $\alpha(K, f)$ and $\beta(K, f)$ will be bounded. As a result, for every $(K, f) \in S(a)$, the total cost converges pointwise to $\alpha(K, f)$ as N goes to infinity. Furthermore, $\alpha(K, f)$ can be computed from (4.52) as:

$$\begin{aligned}\alpha(K, f) \equiv & J_i(K, f, \frac{1}{N}) \mid_{(\frac{1}{N}=0)} = (-K^3a^4c^3\sigma_w^2 - rK^3a^4cf^2\sigma_v^2 + 3K^3a^3bc^3f\sigma_w^2 + rK^3a^3bcf^3\sigma_v^2 \\
& - 3K^3a^2b^2c^3f^2\sigma_w^2 - K^3a^2b^2cf^2\sigma_v^2 - rK^3a^2c^3f^2\sigma_w^2 + rK^3a^2cf^2\sigma_v^2 + K^3ab^3c^3f^3\sigma_w^2 \\
& + K^3ab^3cf^3\sigma_v^2 + rK^3abc^3f^3\sigma_w^2 + rK^3abcf^3\sigma_v^2 + 3K^2a^4c^2\sigma_w^2 + rK^2a^4f^2\sigma_v^2 \\
& - 9K^2a^3bc^2f\sigma_w^2 - rK^2a^3bf^3\sigma_v^2 + 9K^2a^2b^2c^2f^2\sigma_w^2 + K^2a^2b^2f^2\sigma_v^2 + rK^2a^2c^2f^2\sigma_w^2 \\
& - K^2a^2c^2\sigma_w^2 - 2rK^2a^2f^2\sigma_v^2 - 3K^2ab^3c^2f^3\sigma_w^2 - K^2ab^3f^3\sigma_v^2 - rK^2abc^2f^3\sigma_w^2 \\
& + rK^2abf^3\sigma_v^2 + K^2b^2c^2f^2\sigma_w^2 + K^2b^2f^2\sigma_v^2 + rK^2c^2f^2\sigma_w^2 + rK^2f^2\sigma_v^2 - b^2f^2\sigma_w^2 \\
& - 3Ka^4c\sigma_w^2 + 9Ka^3bcf\sigma_w^2 + 3Ka^2c\sigma_w^2 - 9Ka^2b^2cf^2\sigma_w^2 + 3Kab^3cf^3\sigma_w^2 \\
& - 3Kabcf\sigma_w^2 + a^4\sigma_w^2 - 3a^3bf\sigma_w^2 + 3a^2b^2f^2\sigma_w^2 - 2a^2\sigma_w^2 - ab^3f^3\sigma_w^2 + 3abf\sigma_w^2 \\
& + \sigma_w^2)/((a^2(1-cK)^2 - 1)((a-bf)^2 - 1)(1 - a(a-bf)(1-cK))).\end{aligned} \tag{4.58}$$

On the other hand, the isolated (separated) agents cost with $h = 0$, denoted $J_i^{(s)}$, can be similarly calculated. In particular, we have:

$$\begin{aligned}J_i^{(s)}(K, f) &= \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2), \\
&= \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}\left(\begin{bmatrix} x_{k,i} \\ \hat{x}_{k,i}\end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & rf^2\end{bmatrix} \begin{bmatrix} x_{k,i} \\ \hat{x}_{k,i}\end{bmatrix}\right), \\
&= \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1} Tr\left(\begin{bmatrix} 1 & 0 \\ 0 & rf^2\end{bmatrix} \begin{bmatrix} x_{k,i} \\ \hat{x}_{k,i}\end{bmatrix} \begin{bmatrix} x_{k,i} \\ \hat{x}_{k,i}\end{bmatrix}^T\right), \\
&= \overline{\lim_{T\to\infty}} \frac{1}{T}\sum_{k=0}^{T-1} Tr\left(\begin{bmatrix} 1 & 0 \\ 0 & rf^2\end{bmatrix} \bar{P}_s\right), \\
&= Tr\left(\begin{bmatrix} 1 & 0 \\ 0 & rf^2\end{bmatrix} \bar{P}_s\right),\end{aligned} \tag{4.59}$$

where $\bar{P}_s$ is directly calculated from the covariance equation given by

$$A_s \bar{P}_s A_s^T - \bar{P}_s + D_s \Sigma_{w,s} D_s^T = 0, \tag{4.60}$$

with

$$A_s = \begin{bmatrix} a & -bf \\ acK & a(1-cK)-bf \end{bmatrix},\ D_s = \begin{bmatrix} 1 & 0 \\ cK & K \end{bmatrix},\ \Sigma_{w,s} = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}. \tag{4.61}$$

Solving (4.60) and replacing $\bar{P}_s$ in (4.59) yields $J_i^{(s)}(K,f) = J_i(K,f,0)$. In particular, denoting (K^*,f^*) the (K,f) couple minimizing $J_i^{(s)}(K,f)$, we have: $J_i^{(s)}(K^*,f^*) = J_i(K^*,f^*,0)$. Thus the isolated agent optimal policy is asymptotically optimal for the coupled agents provided that $(K^*(a), f^*(a,r)) \in S(a)$. Moreover, if (K,f) is constrained to an arbitrary closed subset M of $S(a)$ containing (K^*,f^*), then invoking the continuity of $\beta(K,f)$ over a closed and bounded set, $\beta(K,f)$ is uniformly bounded on M, and in view of equation (4.57), the convergence of $Argmin_{(K,f)\in M}(J_i(K,f,\frac{1}{N}))$ to (K^*,f^*) will be uniform in M, as N goes to infinity. As a result, over an arbitrary such M, and for N finite, the inequalities given by (4.51) hold. In particular, note that the second inequality is trivial, and the first one can be proven as follows. Since the minimum of $\alpha(K,f)$ is $\alpha(K^*,f^*)$ and noting that the infimum of a sum is greater than or equal to the sum of infimums, then using (4.57) we have

$$\inf_{(K,f)\in S} J_i(K,f,\frac{1}{N}) \geq \alpha(K^*,f^*) + \inf_{(K,f)\in S} \frac{1}{N}\beta(K,f). \tag{4.62}$$

By adding and subtracting $\frac{1}{N}\beta(K^*,f^*)$ to the right hand side of (4.62), we get

$$\inf_{(K,f)\in S} J_i(K,f,\frac{1}{N}) \geq J_i(K^*,f^*,\frac{1}{N}) - \frac{1}{N}\beta(K^*,f^*) + \inf_{(K,f)\in S} \frac{1}{N}\beta(K,f) \tag{4.63}$$

Thus, for N finite, (K^*,f^*) is epsilon optimal where epsilon is $O(\frac{1}{N})$. □

We proceed to study the ϵ-Nash equilibrium property by deriving the dynamics model of a so-called *deviant agent*, say the i_0^{th} agent that tries to improve its cost by choosing a different set of control and estimation gains denoted by $(\tilde{K},\tilde{f})$. In particular, define:

$$m_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} x_{k,j}, \quad \tilde{m}_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} \hat{x}_{k,j}, \tag{4.64}$$

$$\tilde{w}_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} w_{k,j}, \quad \tilde{v}_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} v_{k,j}. \tag{4.65}$$

Then following a procedure similar to that in Section 4.4.1, the 4^{th} order model of the closed-loop deviant agent dynamics can be expressed as:

$$X_{k+1,i_0} = A_d X_{k,i_0} + D_d W_{k,i_0}, \tag{4.66}$$

where the augmented state is

$$X_{k,i_0} = [x_{k,i_0}, \hat{x}_{k,i_0}, m_k^-, \tilde{m}_k^-]^T, \tag{4.67}$$

and matrix A_d is given by

$$A_d = \begin{bmatrix} a & -b\tilde{f} & 0 & 0 \\ a(c+\frac{h}{N})\tilde{K} & a'_{2,2} & ah\tilde{K} & -bhf^*\tilde{K} \\ 0 & 0 & a & -bf^* \\ \frac{(N-1)}{N^2}ahK^* & -\frac{(N-1)}{N^2}bh\tilde{f}K^* & a'_{4,3} & a'_{4,4} \end{bmatrix}, \tag{4.68}$$

with

$$a'_{2,2} = a(1-c\tilde{K}) - b\tilde{f}(1+\frac{h}{N}\tilde{K}), \tag{4.69}$$

$$a'_{4,3} = a(c+\frac{(N-1)}{N}h)K^*, \tag{4.70}$$

$$a'_{4,4} = a(1-cK^*) - bf^*(1+\frac{(N-1)}{N}hK^*), \tag{4.71}$$

and also we have:

$$D_d = \begin{bmatrix} 1 & 0 & 0 & 0 \\ (c+\frac{h}{N})\tilde{K} & h\tilde{K} & \tilde{K} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{(N-1)}{N^2}hK^* & (c+\frac{(N-1)}{N}h)K^* & 0 & K^* \end{bmatrix}. \tag{4.72}$$

Furthermore, the noise vector W_{k,i_0} and its covariance matrix $\Sigma_{w,d}$ are given by:

$$W_{k,i_0} = \begin{bmatrix} w_{k,i_0} \\ \tilde{w}_k^- \\ v_{k+1,i_0} \\ \tilde{v}_{k+1}^- \end{bmatrix}, \Sigma_{w,d} = \begin{bmatrix} \sigma_w^2 & 0 & 0 & 0 \\ 0 & \frac{(N-1)}{N^2}\sigma_w^2 & 0 & 0 \\ 0 & 0 & \sigma_v^2 & 0 \\ 0 & 0 & 0 & \frac{(N-1)}{N^2}\sigma_v^2 \end{bmatrix}. \tag{4.73}$$

The next lemma gives a necessary condition for mean square stability of the deviant agent system.

Lemma 4.5. *A necessary condition for the closed-loop deviant agent dynamics (4.66) to be MSS is that*

$$|a^2(1 - cK^*)(1 - c\tilde{K})(a - bf^*)(a - b\tilde{f})| < 1. \tag{4.74}$$

Proof: Since $\forall N$, $D_d\sqrt{\Sigma_{w,d}}$ is invertible, then the pair $(A_d, D_d\sqrt{\Sigma_{w,d}})$ is controllable, and by virtue of Theorem 3.13, p. 31 in Kumar and Varaiya (1986), the deviant agent system (4.66) will be MSS iff A_d is a stable matrix. Now A_d will be stable iff the Jury stability test is met. In particular, the characteristic polynomial of (4.66) is given by:

$$z^4 + \alpha_3 z^3 + \alpha_2 z^2 + \alpha_1 z + \alpha_0, \tag{4.75}$$

where

$$\alpha_0 = a^2(1 - cK^*)(1 - c\tilde{K})(a - bf^*)(a - b\tilde{f}), \tag{4.76}$$

and the expressions of α_1, α_2, α_3 are complex and are dropped for brevity. Now one of the five stability conditions from the Jury test (see Ogata (1995)) is that $|\alpha_0| < 1$, which yields (4.74). $\square$

Theorem 4.5. *The set of gains $(K^*(a), f^*(a, r)) \in S(a)$ defines an ϵ-Nash equilibrium.*

Proof: The deviant agent must always opt for a $(\tilde{K}, \tilde{f})$ couple which would stabilize matrix A_d (Hurwitz matrix), no matter what N is. Note that the set of such stabilizing couples is non empty since it includes (K^*, f^*) irrespective of N. Also, note that in view of (4.74), $\tilde{K}$ can become unbounded only if $a = b\tilde{f}$. Additionally, $\tilde{f}$ can be unbounded only if $c\tilde{K} = 1$. Let us refer to these cases as (i) and (ii), respectively. We now establish that both (i) and (ii) are excluded if the deviant agent optimizes its choices. Indeed case (i) would imply from (4.4) that $\overline{\lim}_{k\to\infty}\hat{x}_{k,i_0}$ can only remain of bounded variance, if y_{k,i_0} goes to zero almost everywhere as $k \to \infty$. However, this is impossible in view of the independent measurement noise that enters the i_0^{th} agent observations. Case (ii) on the other hand, would imply from (4.3) that $\overline{\lim}_{k\to\infty}\hat{x}_{k,i_0} = 0$ a.e., otherwise $\overline{\lim}_{k\to\infty}\mathbb{E}[u_{k,i_0}^2] \to \infty$; which would be clearly suboptimal since f^*, K^* achieve a finite limiting cost. However, this cannot happen since from (4.4) and $\tilde{K} \neq 0$, this would mean $\overline{\lim}_{k\to\infty}y_{k,i_0} = 0$ a.e., and we know the latter to be impossible. As a result for any $a \in S^{opt}(a, r)$, where $S^{opt}(a, r)$ denotes the set of couples $\{(a, r)|(K^*(a), f^*(a, r)) \in S(a)\}$, $\tilde{K}(N)$ and $\tilde{f}(N)$ will be bounded. Thus, if the deviant agent maintains its choice fixed of $\tilde{K}(N_0)$ and $\tilde{f}(N_0)$, for any $N \geq N_0$, then this choice becomes eventually suboptimal for N large enough. Indeed, as N increases to infinity, the dynamics of $[x_{k,i_0}, \hat{x}_{k,i_0}]^T$ becomes entirely decoupled from that of $[m_k^-, \tilde{m}_k^-]^T$, as $\frac{(N-1)}{N^2} x_{k,i_0}$, $\frac{(N-1)}{N^2} \tilde{f}(N_0)\hat{x}_{k,i_0}$, $\tilde{K}(N_0)m_k^-$, $\tilde{K}(N_0)\tilde{m}_k^-$ go to zero almost surely. In that case, the deviant agent optimal choice for $(\tilde{K}(N), \tilde{f}(N))$ becomes precisely (K^*, f^*). As a result, any cost improvement attributed to some optimal decision $(\tilde{K}^*(N_0), \tilde{f}^*(N_0))$ by the deviant agent can only improve its cost by $\epsilon(N)$, with $\epsilon(N)$ going to zero as $N \to \infty$. This establishes that a finite N_0-based policy, is such that the (K^*, f^*) policy is an ϵ-Nash equilibrium. Finally, note that if the deviant agent starts at the outset with the $N = \infty$ policy, from Theorem 4.1, it would have to pick the (K^*, f^*) policy as an optimal response. $\square$

Numerical results IV

In this section, we present a numerical example on the deviant agent (namely, the 50^{th} agent) in a population of $N = 100$ agents via simulations of $T = 100000$ steps. In particular, we consider $a = 2.44$, where the optimal Kalman-Riccati couple $(K^* = 0.8595, a_{f^*} = 0.3428)$ belongs to the stability region and all the agents except the deviant agent, apply this couple. We let the deviant agent apply another stabilizing couple $(\tilde{K} = 0.8307, a_{\tilde{f}} = 0.3929)$ such that all eigenvalues of A_d lie inside the unit circle. The improvement on the cost achieved by the deviant agent $i_0 = 50$ is shown in Figure 4.6. However, performing the same simulation with $N = 1000$ agents, we observed that the cost difference has disappeared.

Boundaries of stability regions of separated optimal gains

Assuming $a > 0$ for the sake of analysis in the whole forthcoming discussion, we first define a threshold on a.

Definition 4.1. *a_{Nash} is the (unique) value of a if it exists, such that:*

$$a_{Nash} - bf_{\sup}(a_{Nash}) = 0. \tag{4.77}$$

Also, we have the following Lemma.

Lemma 4.6. *Assume Eq. (4.77) admits a solution a_{Nash}, and let $f_{\sup}(a)$ be given by (4.46). Then we have:*

(i) a_{Nash} exists uniquely;

(ii) for $a > a_{Nash}$, $a/b > f_{\sup}(a)$;

(iii) for $a < a_{Nash}$, $a/b < f_{\sup}(a)$;

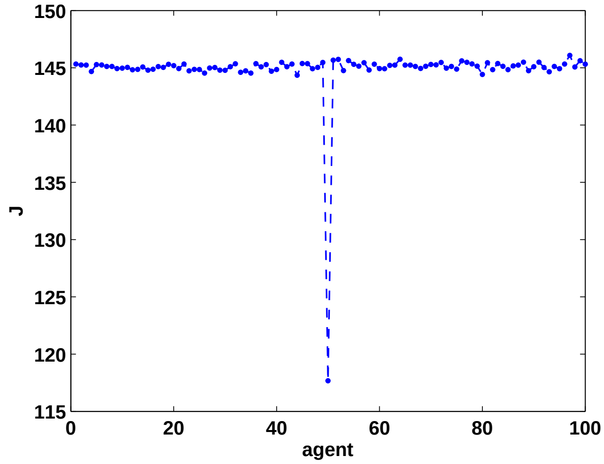

Figure 4.6 Cost comparison between the deviant agent and each of the other agents when $a = 2.44$, under $r = 1$, $N = 100$, $i_0 = 50$.

(iv) a_{Nash} can assume any positive value by properly modifying h.

Proof: First of all, from the expression of $K^*(a)$ (see Algorithm 1), it is straightforward to see that $K^*(a) > 0$ for all a and that $cK^*(a) \in (0, 1)$ for all a. A direct computation of the derivative of $K^*(a)$ allows one to also verify that $K^*(a)$ is a strictly increasing function of $a > 0$ which ranges from $\frac{1}{c(1+\delta)}$ when $a = 0$ (where $\delta \triangleq \sigma_v^2/(c^2\sigma_w^2)$ is a positive quantity) and goes toward $1/c$ as $a \to \infty$.

Consider now the difference $f_{\sup}(a) - a/b$, which can be written as follows:

$$f_{\sup}(a) - a/b = \frac{ag(a,h)}{b[1 + a(1 - cK^*(a)) + hK^*(a)]}\,, \tag{4.78}$$

where

$$g(a,h) = 1 - (c + h)K^*(a) + 1/a. \tag{4.79}$$

Notice that from (4.36) the denominator of (4.78) is positive for all $a > 0$, so the sign of $f_{\sup}(a) - a/b$ coincides with the sign of $g(a,h)$ and $f_{\sup}(a) = a/b$ if and only if $g(a,h) = 0$. Given the aforementioned properties on $K^*(a)$, we have that $g(a,h)$ is strictly decreasing for $a > 0$ and ranges from $+\infty$ (when $a \to 0$) to $-h/c$ (when $a \to \infty$). Therefore, by continuity, $g(a,h)$ will cross zero at some $a = a_{Nash}$, and by the strict decreasing character of $g(a,h)$, the intersection will be unique. This concludes the proof of (i) and also of (ii) and (iii) since, as mentioned, the sign of $f_{\sup}(a) - a/b$ coincides with the sign of $g(a,h)$.

Property (iv) can be obtained by considering, for $h_2 > h_1$, the difference:

$$g(a,h_2) - g(a,h_1) = (h_1 - h_2)K^*(a) < 0. \tag{4.80}$$

In particular:

$$g(a_{Nash_1}, h_2) - g(a_{Nash_1}, h_1) = g(a_{Nash_1}, h_2) < 0, \tag{4.81}$$

which indicates: $a_{Nash_1} > a_{Nash_2}$, that is, the solution $a_{Nash}(h)$ of (4.77) is strictly decreasing with h. Moreover, as $h \to 0$, we have $a_{Nash} \to +\infty$ and, as $h \to +\infty$, since $f_{\sup}(a) \to 0$ then $a_{Nash} \to 0$. Thus, by moving h from zero to infinity, one moves a_{Nash} over its complete possible range. $\square$

Remark 4.3. *Noting the expression of $g(a, h)$ given by (4.79), it is possible to evaluate the threshold h_t such that a_{Nash} crosses 1 (i.e. $a_{Nash} > 1$ for all $h < h_t$ and $a_{Nash} < 1$ for all $h > h_t$) which is $h_t = (2 - cK^*(1))/K^*(1)$.*

The threshold a_{Nash} represents the maximum value of a past which it is not always possible to apply so-called isolated optimal control policies without causing potential instability problems. More specifically, if $a < a_{Nash}$, we have that $(K^*(a), f^*(a, r))$ stabilizes the population for all r while, for $a > a_{Nash}$, the range of r progressively decreases. The following proposition provides a formal characterization of this behavior.

Proposition 4.1. *Assume $a > 0$, let $\bar{f}_{sup}(a) = \min\{a/b, f_{\sup}(a)\}$ and define*

$$r_{\inf}(a) = \frac{-b(a - b\bar{f}_{sup}(a))}{\bar{f}_{sup}(a)[a(a - b\bar{f}_{sup}(a)) - 1]}. \tag{4.82}$$

Then there exists a threshold $a_{\sup} > \max\{a_{Nash}, 1\}$, which satisfies

$$a_{\sup}[a_{\sup} - bf_{\sup}(a_{\sup})] - 1 = 0, \tag{4.83}$$

such that, for all $a < a_{\sup}$, $(K^*(a), f^*(a, r)) \in S(a)$ for all $r > r_{\inf}(a) \geq 0$ (with $r_{\inf}(a) = 0$ for all $a \leq a_{Nash}$). Moreover, $r_{\inf}(a) \to +\infty$ as $a \to a_{\sup}$, and, if $a > a_{\sup}$, $(K^*(a), f^*(a, r)) \notin S(a)$ for all r.

Proof: See the Appendix. $\square$

Remark 4.4. *According to Proposition 4.1 and Lemma 4.6, for all $0 < a < a_{Nash}$, $(K^*(a), f^*(a, r)) \in S(a)$ for all positive r while, for $a_{Nash} < a < a_{\sup}$, $(K^*(a), f^*(a, r)) \in S(a)$ only if r is larger than a positive threshold $r_{\inf}(a)$. Moreover, since $r(\bar{f}_{inf}(a)) = +\infty$, for $a = a_{\sup}$, $r_{\inf}(a)$ reaches $+\infty$. So, at that point, the range of admissible r's shrinks to zero. Past $a_{\sup}$, isolated interference indifferent optimal controls can no longer stabilize the system for any value of r, and one has to resort to cooperative control.*

Numerical results V

Figure 4.7 shows a numerical investigation on $r_{\inf}(a)$. More specifically, considering values of a ranging on a grid with step 0.01 one can observe that:

— For $0 < a \leq 1.79$, we have $r_{\inf}(a) = 0$;

— For $1.8 \leq a \leq 2.53$, $r_{\inf} > 0$ and rapidly goes toward infinity;

— For $a = 0$ and $a \geq 2.54$, reverse engineering does not apply for any r, so $r_{\inf}$ is not defined.

Thus, we have: $a_{Nash} = 1.79$, and $a_{\sup} = 2.53$. We note that the analytical values of a_{Nash} and $a_{\sup}$ obtained through (4.77) and (4.83), i.e., respectively 1.7963 and 2.5369, confirm the numerical findings.

4.6 Cooperative decentralized separated policies

Definition 4.2. *Cooperative decentralized separated policies are defined as decentralized separated policies (see Section 4.3) with common gains K, f such that the resulting social cost*

$$J_{soc}^{(N)} = \frac{1}{N}\sum_{j=1}^{N} J_j \tag{4.84}$$

is minimized.

If $a > a_{\sup}$, then agents must cooperate for otherwise, they risk having to pay an infinite cost. This is a situation where the optimal Kalman-Riccati couple (K^*, f^*) is outside of the stability region. On the other hand, even when $a \leq a_{\sup}$, agents may still be interested in achieving optimal *cooperative* decentralized separated policies. We have the following lemma and proposition for the cooperative cost.

Lemma 4.7. *(K, f) based local control policies, $\forall (K, f) \in S(a)$, when uniformly applied by all agents, for a given N lead to a steady-state cost given by:*

$$J_i = Tr(\bar{Q}\bar{P}_\infty) \tag{4.85}$$

where $\bar{Q} = diag([1, rf^2, 0, 0])$, and $\bar{P}_\infty$ is the steady-state solution of the covariance equation given by

$$\bar{P}_{k+1} = A\bar{P}_k A^T + D\Sigma_w D^T \tag{4.86}$$

Proof: See the cost calculations in Theorem 4.4. □

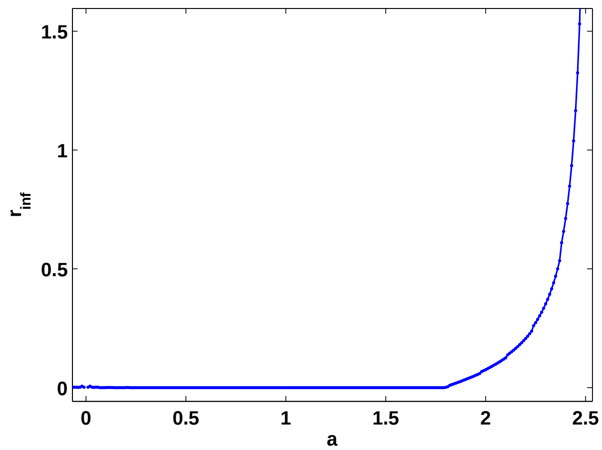

Figure 4.7 $r_{\text{inf}}(a)$ as calculated using increments of 0.01 by evaluating for values of a ranging on a grid of step 0.01.

Proposition 4.2. *If the optimal Kalman-Riccati couple (K^*, f^*) belongs to stability region $S(a)$, then it characterizes an ϵ-optimal cooperative decentralized separated control policy for system (4.1), (4.2), (4.5), with ϵ of order $1/N$, over any closed subset of $S(a)$ including (K^*, f^*).*

Proof: The proof follows readily by recalling individual cost expression (4.57), and recognizing that (i) $\beta(K, f)$ is continuous in (K, f) over a closed subset of $S(a)$, and therefore is uniformly bounded over that subset, (ii) the minimum of $\alpha(K, f)$ is $\alpha(K^*, f^*)$. $\square$

In the following, we numerically explore the situation when (K^*, f^*) does not belong to the stability region $S(a)$.

4.6.1 Numerical results VI

In this section, we present some numerical results on cooperative control. Figs. 4.8-4.12 show the simulation results for $a = 2.44$, where $(K^* = 0.8595, a_{f^*} = 0.3428)$ belongs to the stability region. It is observed that the optimal cooperative $(K = 0.8307, a_f = 0.3929)$ is near the Kalman-Riccati couple $(K^* = 0.8595, a_{f^*} = 0.3428)$ and approaches this point as N goes to infinity.

Moreover, the simulation results for $a = 4$ are depicted via Figs. 4.13-4.14. We note that the reverse engineering of r does not apply for $a = 4$, and that K and a_f do not converge toward K^* and a_{f^*} as N increases. Also in this case, $(K^* = 0.9414, a_{f^*} = 0.2344)$ does not belong to the stability region.

Furthermore, for $a = 4$, we present a numerical example on the deviant agent (i.e., the 50^{th} agent) in a population of $N = 100$ agents via simulations of $T = 100000$ steps. In

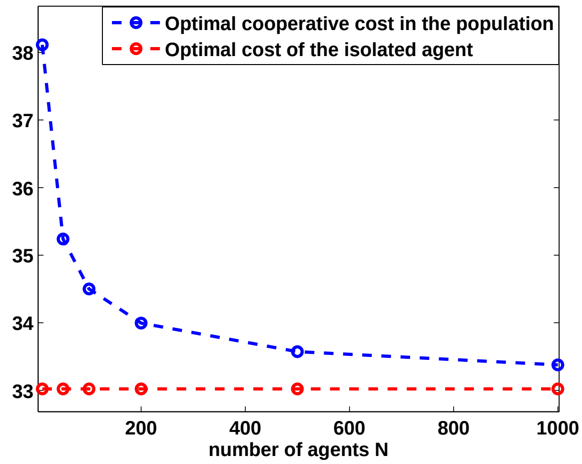

Figure 4.8 Optimal cost J_i^* when $a = 2.44$, under $r = 1$, $b = c = h = 1$, $\bar{x}_0 = 0$, $\sigma_0 = \sigma_w = \sigma_v = 1$.

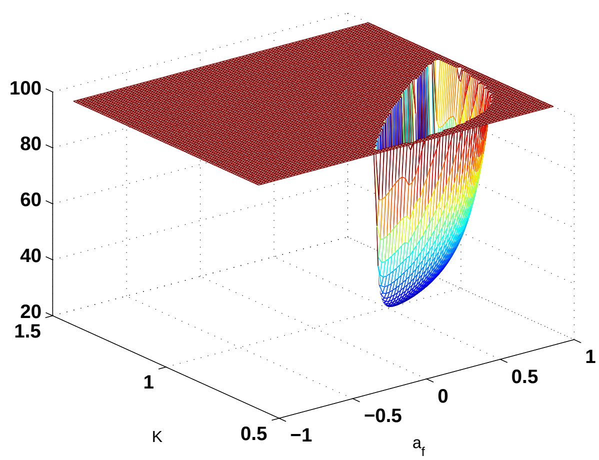

Figure 4.9 Analytical cost J_i as a function of K and a_f, saturated at 100, where $a_f = a - bf$, $a = 2.44$, $N = 100$.

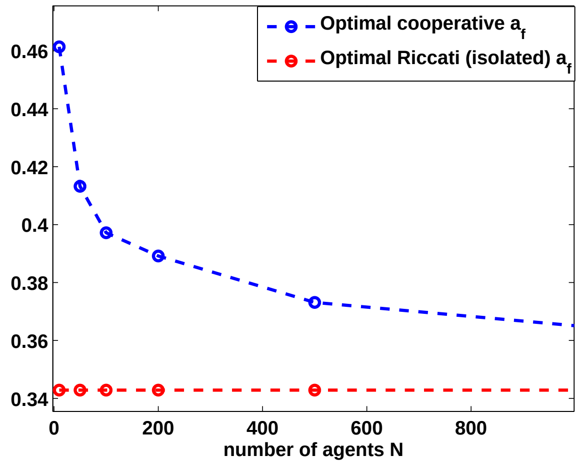

Figure 4.10 Optimal cooperative a_f as a function of the number of agents, where $K = K^*$ and $a = 2.44$.

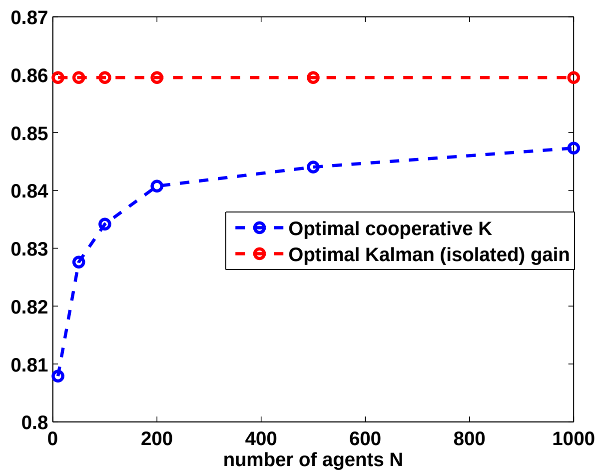

Figure 4.11 Optimal cooperative K as a function of the number of agents, where $f = f^*$ and $a = 2.44$.

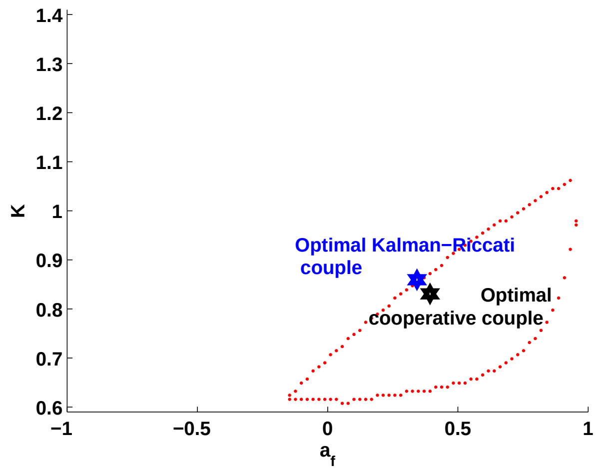

Figure 4.12 Stability region (inside the red border) when $a = 2.44$, $N = 100$.

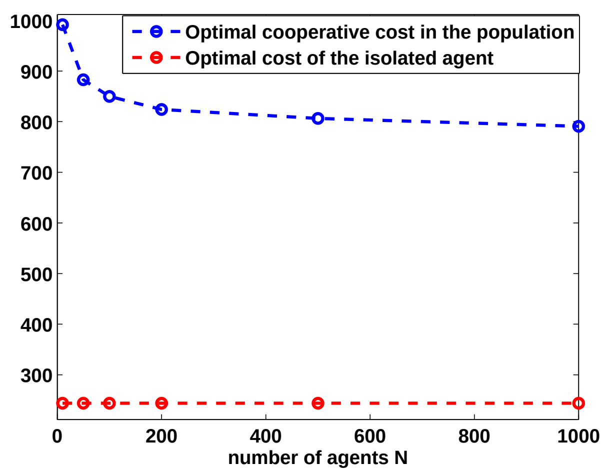

Figure 4.13 Optimal (analytical) cost J_i^* when $a = 4$.

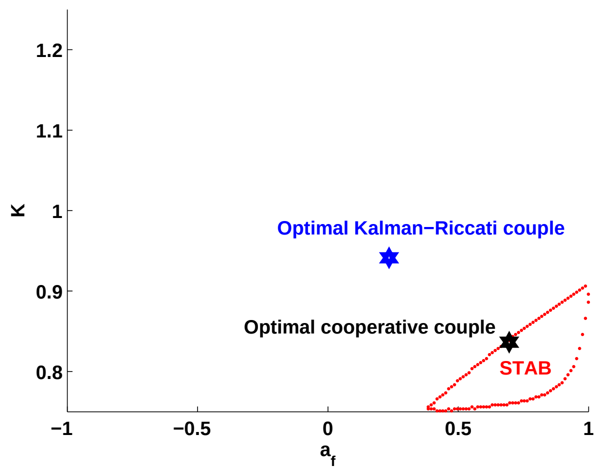

Figure 4.14 Stability region (inside the red border) when $a = 4$, $N = 100$.

particular, all the agents except the deviant agent, apply the optimal cooperative couple ($K = 0.8365, a_f = 0.6955$). We let the deviant agent apply another stabilizing couple ($\tilde{K} = 0.8842, a_{\tilde{f}} = 0.4857$) such that all eigenvalues of A_d lie inside the unit circle. The improvement on the cost achieved by the deviant agent $i_0 = 50$ is shown in Figure 4.15.

We note that there does not exist an optimum cooperative control setting in the $a > a_{\text{sup}}$ region, but an infimum of the cost which would be its minimal value on the stability border under the assumption that $J_i(K, f)$ is strictly convex in K and f. For example, when $a = 3$, we have performed long simulation runs to evaluate the isolated cost on a grid of the (K, a_f) plane with a step of about 0.01 to determine the minimum. The population stability region (inside the red dot points) is depicted in Figure 4.16, where we verify that the minimum of the isolated cost on such stability region falls indeed on its border. It also shows the cost of N agents as N increases, and the position of these minima. It is observed that the optimal $(K(N), a_f(N))$ couple is actually approaching the optimal isolated cost on the border as N increases. Moreover, considering the black segment in this figure (called segment Σ) which comprises 100 points numbered from inside to the border of the stability region, Figure 4.17 illustrates the isolated cost versus the cost of N agents for different values of N. Furthermore, we study the convergence rate with respect to N when approaching the stability border. In particular, consider a couple of gains (K, f); let $\epsilon > 0$ be a small positive number, and also let $N_{\min}$ be the number of agents such that $|J^{(S)}(K, f) - J^{(N)}(K, f)| < \epsilon$ for all $N > N_{\min}$, where under (K, f), $J^{(N)}$ denotes the cost of each agent in a population of N elements and $J^{(S)}$ is the isolated cost (see 4.59). As shown in Figure 4.18, it is observed that the minimum value of N such that $|J^{(S)}(K, f) - J^{(N)}(K, f)| < \epsilon$ is rapidly increasing as one approaches the stability border. In essence, these numerical results confirm: (i) the infimum cooperative cost

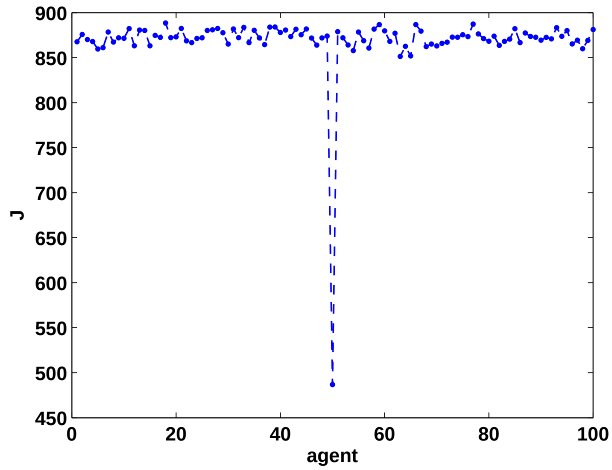

Figure 4.15 Cost comparison between the deviant agent ($i_0 = 50$) and each of the other agents when $a = 4$, under $r = 1$, $N = 100$.

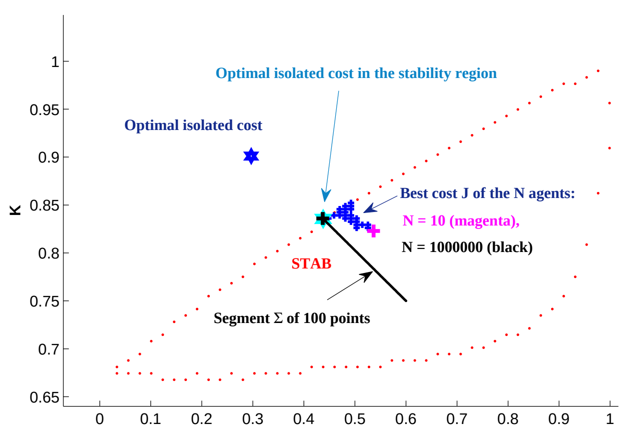

Figure 4.16 Stability region (inside the red border), and the best cost J of the N agents with increasing N when $a = 3$; the magenta and black colors correspond to $N = 10$ and $N = 1000000$, respectively.

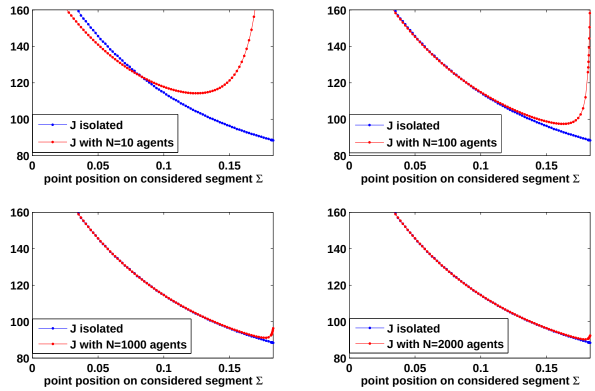

Figure 4.17 Isolated cost and cost of each agent for different values of N on the points of segment Σ.

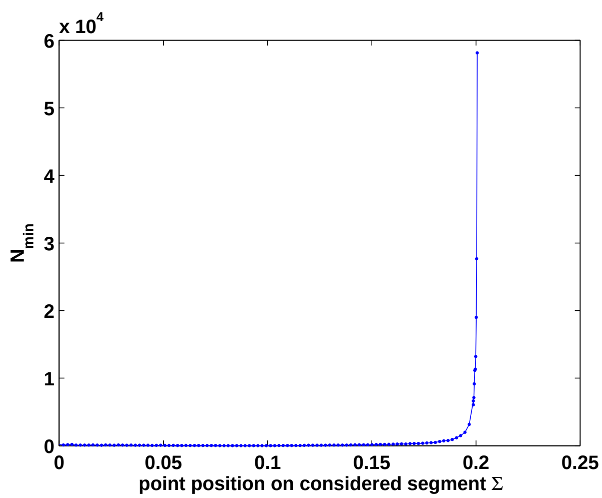

Figure 4.18 Minimum number of agents N such that $|J^{(S)} - J^{(N)}| < \epsilon$ (with $\epsilon = 1$), along segment Σ.

is on the stability border; (ii) the slowing down of convergence of actual cost with N agents to the isolated cost within the stability region, when N goes to infinity, as one approaches the edge of the stability region.

4.7 Application to wireless communications

In this section, we present an application example for decentralized power control in code division multiple access (CDMA) cellular telephone systems (Aziz and Caines (2014); Huang *et al.* (2004); Koskie and Gajic (2006); Perreau and Anderson (2006)).

Following Tse and Hanly (1999); Verdú and Shamai (1999), we consider a mobile system network in the context of a large number of users with a signal processing gain assumed to be proportional to $1/N$. Let $p_{k,i}^{(m)}$ and $\alpha_{k,i}$ denote, respectively, the transmitted power and the mean squared value of the uplink channel gain for the i^{th} mobile user of the network and let $p_{k,i}^{(b)}$ denote the received power at the base station for user i, where $p_{k,i}^{(b)} = \alpha_{k,i} p_{k,i}^{(m)}$. Based on the work in Perreau and Anderson (2006), we model the received power dynamics at the base station by

$$p_{k+1,i}^{(b)} = p_{k,i}^{(b)} + u_{k,i} + w_{k,i} \tag{4.87}$$

with observations given by:

$$y_{k,i} = p_{k,i}^{(b)} + \frac{h}{N} \sum_{j \neq i}^{N} p_{k,j}^{(b)} + \sigma_{th}^2 + v_{k,i}, \tag{4.88}$$

which is the average over slot k of the power of the CDMA signal despread by the spreading sequence of user i, where σ_{th}^2 is the variance of the background thermal noise process (modeled as a zero mean Gaussian random variable). Note that the resulting signal processing gain is assumed to be h/N.

The goal is to perform decentralized power control in order to design the control command $d_{dB_{k,i}}$ which updates the transmitted power (on the logarithmic scale, i.e., $d_{dB_{k,i}} = 10\log_{10}(d_{k,i})$, for its ease of implementation in practical systems) according to

$$p_{dB_{k+1,i}}^{(m)} = p_{dB_{k,i}}^{(m)} + d_{dB_{k,i}}, \tag{4.89}$$

so that the signal to interference plus noise ratio of each user achieves a target value γ (common to all users), i.e., $SINR_{k,i} = \gamma$. This is feasible at minimum power if the received powers for all users are equal to $\bar{p}^*$ given by Perreau and Anderson (2006)

$$\bar{p}^* = \frac{\gamma \sigma_{th}^2}{1 - \gamma h (N-1)/N} \,. \tag{4.90}$$

Therefore, we minimize the individual cost function for each user as defined by

$$J_i \triangleq \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}\{(p_{k,i}^{(b)} - \bar{p}^*)^2 + ru_{k,i}^2\}\cdot \tag{4.91}$$

By applying change of variables

$$x_{k,i} = p_{k,i}^{(b)} - \bar{p}^*, \quad y'_{k,i} = y_{k,i} - (1 + h - \frac{h}{N})\bar{p}^* \tag{4.92}$$

to (4.87), (4.88) and (4.91), we obtain a system in the form of (4.1), (4.2) and (4.5) with $a = b = 1$, $c = 1 - \frac{h}{N}$ and $v'_{k,i} = \sigma_{th}^2 + v_{k,i}$. Note that we formulated this problem on the linear scale and that it is trivial to derive a relative change in watts into a change in decibels.

4.7.1 Numerical results VII

In this section, we consider a network with $N = 100$ users which are assumed to be equally spaced on a circle around the base station. Choosing $\gamma = 0.95$, $\sigma_{th}^2 = 1$ and $h = 1$ in (4.90) yields $\bar{p}^* = 15.9664$. We apply Algorithm 1 with $a = b = 1$, $c = 0.99$ and the pair $(K^* = 0.5017, a_f = 0.8)$ within the stability region. Reverse engineering the cost structure (4.5) for parameter $r > 0$, we get $r = 20$. Figure 4.19 shows the power tracking error $x_{k,i} = p_{k,i}^{(b)} - \bar{p}^*$ for a representative user (namely, the 50^{th} mobile user). Moreover, Figure 4.20 illustrates the transmitted power $p_{dB_{k,50}}^{(m)}$ with $\alpha_{50} = 0.7$.

4.8 Conclusion

We have studied a class of interference induced games in a system of uniform agents coupled via their distinct sets of partial observations, whereby each agent has noisy measurements of its own state. We have shown that interference coupled agents can afford to act non cooperatively provided their individual stability level as characterized by a quantity called a_{Nash}, is sufficient relative to the signal to noise ratio in their observations and the number of agents is sufficiently high. Moreover, there is a lack of stability threshold past which, the only choice left for the majority of agents is to act cooperatively. The apparent role of individual agent systems lack of stability in interference coupled systems, points at a potential hitherto unsuspected role of instability in more classical mean field games where the mean agent state enters individual dynamics (Huang *et al.* (2005)).

In future work, we will investigate the use of optimal (growing dimension) estimators to improve the situation. We will also generalize our analysis to the multivariate case. Finally, we note that the finite number of agents version of the game constitutes a challenging problem

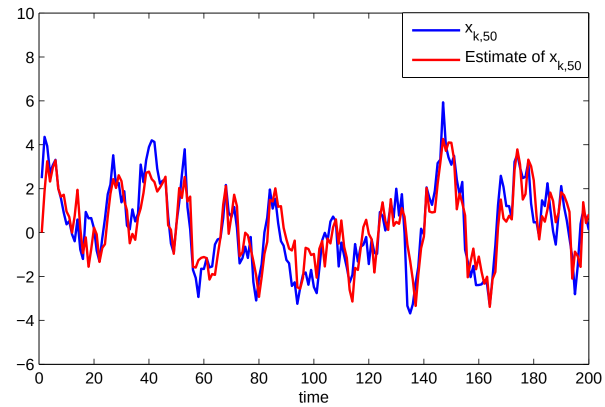

Figure 4.19 The power tracking error $x_{k,i} = p_{k,i}^{(b)} - \bar{p}^*$ and its estimate when $r = 20$, $\bar{p}^* = 15.9664$, $N = 100$, $i = 50$.

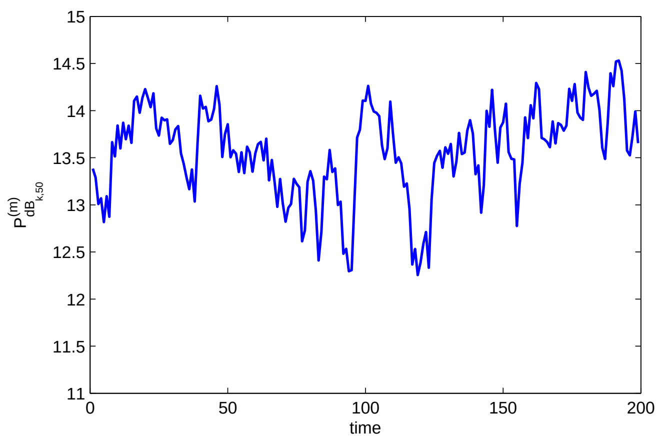

Figure 4.20 Transmitted power $p_{dB_{k,i}}^{(m)}$ in decibels when $\alpha_i = 0.7$, $i = 50$, in a group of 100 mobile users.

in signalling and decentralization.

4.9 Appendix

Proof of Proposition 4.1. First of all, $f^*(a,r) = ab\Sigma/(b^2\Sigma + r)$, where Σ is the positive solution of (4.49). It is straightforward to see that, with $a > 0$, we have $f^*(a,r) > 0$ and $f^*(a,r) < a/b$. This implies also that, if $f^*(a,r)$ is stabilizing, then $0 < a - bf^*(a,r) < 1$.

Based on (4.47), the fact that $r(f)$ must be positive and taking into account that $a - bf^*(a,r) > 0$, we have that the possible range of $f^*(a,r)$ is $(0, +\infty)$ if $a \in (0,1]$, and $((a^2-1)/(ab), +\infty)$ if $a > 1$.

Finally, if $f^*(a,r)$ is stabilizing the population, then $f^*(a,r) \in (f_{inf}(a), f_{\sup}(a))$, where it is easy to verify that: $f_{inf}(a) \leq 0$ if $a \in (0,1)$ (since $f = 0$ is stabilizing given that with $f = 0$ and $a \in (0,1)$, Eqs. (4.31)-(4.34) are satisfied) and that, according to (4.44), $f_{inf}(a) = (a-1)/b < (a^2-1)/(ab)$ if $a > 1$.

Putting together all the above constraints, we can conclude that $(K^*(a), f^*(a,r)) \in S(a)$ implies that $f^*(a,r) \in (\bar{f}_{inf}(a), \bar{f}_{sup}(a))$, where $\bar{f}_{inf}(a) = 0$ if $a \in (0,1]$ and $\bar{f}_{inf}(a) = (a^2-1)/(ab)$ if $a \geq 1$. Notice that $f_{inf}(a) \leq \bar{f}_{inf}(a)$ for all a.

Now, for all $f \in (\bar{f}_{inf}(a), \bar{f}_{sup}(a))$, the derivative of $r(f)$ with respect to f which is given by:

$$\frac{\partial r}{\partial f} = \frac{ab[(a-bf)^2 - 1]}{f^2[a(a-bf)-1]^2}, \tag{4.93}$$

is negative. This happens as a result of the fact that the interval $(\bar{f}_{inf}(a), \bar{f}_{sup}(a))$ is included in the stability interval $(f_{inf}(a), f_{\sup}(a))$ where $(a-bf)^2 < 1$.

Notice also that $r(\bar{f}_{inf}(a)) = +\infty$. So the values of $r(f)$, for f ranging in $(\bar{f}_{inf}(a), \bar{f}_{sup}(a))$, will decrease from $+\infty$ down to the value $r_{\inf}(a)$ given in (4.82). Furthermore, there is a one to one map between f and the corresponding r.

Finally, in view of Lemma 4.6, for $a \in (0, a_{Nash})$, individuals applying optimal isolated policies will stabilize for the whole range of r from 0 to infinity (since in this case $\bar{f}_{sup}(a) = a/b$ and hence $r_{\inf}(a) = 0$). For a past a_{Nash}, $r_{\inf}(a)$ is strictly positive, since $\bar{f}_{sup}(a) = f_{\sup}(a) < a/b$ where r is still positive.

As a keeps on increasing past a_{Nash}, at some point we have $\bar{f}_{sup}(a) \equiv \bar{f}_{inf}(a)$. This can occur only if $a > 1$, since, for $a \in (0,1]$, $\bar{f}_{sup}(a) > 0$ and $\bar{f}_{inf}(a) \leq 0$. On the other hand, if $a > 1$, $\bar{f}_{sup}(a) = \bar{f}_{inf}(a)$ if and only if a is such that $(c+h)K^*(a) - 1 = \frac{1}{a}(hK^*(a)+1)$. Following a procedure similar to the one adopted in the proof of Lemma 4.6, it is possible to show that the value of $a > 1$ such that $(c+h)K^*(a) - 1 = \frac{1}{a}(hK^*(a)+1)$ uniquely exists and that, denoting $a_{\sup}$ this value, we have $a_{\sup} > a_{Nash}$.

Notice also that, for $a > a_{Nash}$, $\bar{f}_{sup}(a) = f_{\sup}(a)$ and that, for $a > 1$, $\bar{f}_{inf}(a) = (a^2 -$

$1)/(ab)$. For this reason, a_{sup} is the value of a such that $f_{\text{sup}}(a) = (a^2-1)/(ab)$ which yields (4.83).

For $a > a_{\text{sup}}$, $\bar{f}_{sup}(a) = f_{\text{sup}}(a) < (a^2-1)/(ab) = \bar{f}_{inf}(a)$, i.e., $(K^*(a), f^*(a,r)) \notin S(a)$ for all r. This concludes the proof. $\square$

CHAPTER 5 ARTICLE 2 : FILTERING FOR DECENTRALIZED CONTROL IN MULTI-AGENT INTERFERENCE COUPLED SYSTEMS

Mehdi Abedinpour Fallah, Roland P. Malhamé, David Saussié and Francesco Martinelli
Submitted to *IEEE Transactions on Control Systems Technology*, January 2017

5.1 Abstract

The object of study in the recent theory of Mean Field Games has been primarily large populations of agents interacting through a population dependent coupling term, entering through individual cost or dynamics. However, there are situations where agents are essentially independent, except for measurement interference. This is the case for example in cellular communications networked control across noisy channels.

In previous work, we formulated the case of interference coupled linear partially observed stochastic agents as a game. Conditions were developed under which naively ignoring the interference term leads to asymptotically (in population size) optimal control laws which are Riccati gain based. We tackle here the case of exact decentralized filtering under a class of time invariant certainty equivalent feedback controllers, and numerically investigate both stabilization ability and performance of such controllers as the state estimate feedback gain varies. While the optimum filters have memory requirements which become infinite over time, the stabilization ability of their finite memory approximation is also tested.

5.2 Introduction

Large population stochastic multi-agent systems have gained significant attention in the control community in recent years. This is due to the rich theory associated with decentralized control and system performance as well as to the growing number of important and challenging applications in control of networked dynamical systems, such as wireless sensor networks (Chong and Kumar (2003)), very large scale robotics (Reif and Wang (1999)), controlled charging of a large population of electric vehicles (Karfopoulos and Hatziargyriou (2013)), synchronization of coupled oscillators (Yin *et al.* (2012)), swarm and flocking phenomenon in biological systems (Grönbaum and Okubo (1994); Passino (2002)), evacuation of large crowds in emergency situations (Helbing *et al.* (2000); Lachapelle (2010)), sharing and competing for resources on the Internet (Altman *et al.* (2006)), to cite a few. It is common in multi-agent systems to limit each individual agent in the system in terms of what it can decide on its own, what it can do on its own, and what it can measure on its own about its local environment. Owing to the limited sensing ability, it is not feasible for each individual agent to collect all other agents' state information, especially for large-scale dynamic systems. Therefore, the design of decentralized control and estimation laws depending only on local state measurements are required.

Several decentralized and distributed estimation schemes for large-scale systems have been proposed to make the estimation problem computationally efficient. In Roshany-Yamchi *et al.* (2013); Ruess *et al.* (2011) distributed and decentralized approaches to state estimation and control were developed for large-scale multi-rate systems with applications to power networks and plantwide processes, respectively. In Caines and Kizilkale (2013, 2014, 2016); Huang *et al.* (2006a); Wang and Zhang (2013) distributed decision-making with partial observation for large population stochastic multi-agent systems was studied, where the synthesis of Nash strategies was investigated for the agents that are weakly coupled through either individual dynamics or costs.

In previous work Abedinpour Fallah *et al.* (2016), the case of N uniform agents described by linear stochastic dynamics with quadratic costs and partial linear observations involving the mean of all agents was considered, and the problem was formulated as an interference induced game. We explored conditions under which a Luenberger like observer, together with a constant state feedback in individual systems would be : (i) ideally optimal, (ii) at least stable. Our objective in the current paper is to extend the class of candidate stabilizing control structures via optimal filtering. In particular, we study the optimal decentralized filtering problem under a class of certainty equivalent controllers, and numerically investigate both stabilization ability and performance of such controllers as the state estimate feedback gain varies. While the optimum filters have memory requirements which become infinite over time, the stabilization ability of their finite memory approximation is also tested.

The rest of the paper is organized as follows. The problem is defined and formulated in Section 5.3. A summary of the previous work Abedinpour Fallah *et al.* (2016) is given at the start of Section 5.4, followed by detailed derivations of the optimal growing dimension filter and associated approximate finite memory filters. In Section 5.5, both stabilization ability and performance of this class of state estimate feedback controllers are numerically investigated. Section 5.6 presents an application example for a class of networked multi-agent control systems. Concluding remarks are stated in Section 5.7.

5.3 Model Formulation and Problem Statement

Consider a system of N agents, with individual scalar dynamics. The evolution of the state component is described by

$$x_{k+1,i} = ax_{k,i} + bu_{k,i} + w_{k,i}, \tag{5.1}$$

with partial scalar state observations given by

$$y_{k,i} = cx_{k,i} + h\left(\frac{1}{N}\sum_{j=1}^{N} x_{k,j}\right) + v_{k,i}, \tag{5.2}$$

for $k \geq 0$ and $1 \leq i \leq N$, where $x_{k,i}, u_{k,i}, y_{k,i} \in \mathbb{R}$ are the state, the control input and the measured output of the i^{th} agent, respectively. The random variables $w_{k,i} \sim \mathcal{N}(0, \sigma_w^2)$ and $v_{k,i} \sim \mathcal{N}(0, \sigma_v^2)$ represent independent Gaussian white noises at different times k and at different agents i. The Gaussian initial conditions $x_{0,i} \sim \mathcal{N}(\bar{x}_0, \sigma_0^2)$ are mutually independent and are also independent of $\{w_{k,i}, v_{k,i}, 1 \leq i \leq N,\ k \geq 0\}$. σ_w^2, σ_v^2 and σ_0^2 denote the variance of $w_{k,i}$, $v_{k,i}$ and $x_{0,i}$, respectively. Moreover, a is a scalar parameter and $b, c, h > 0$ are positive scalar parameters. In addition, the individual cost function for each agent is given by

$$J_i \triangleq \overline{\lim_{T\to\infty}} \frac{1}{T}\mathbb{E}\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2). \tag{5.3}$$

where $r > 0$ is a positive scalar parameter, $\mathbb{E}\{.\}$ is the expectation operator, and $\overline{\lim}$ is $\limsup$.

The problem to be considered is to develop decentralized estimation policies such that each agent is stabilized by a linear feedback control of the form

$$u_{k,i} = -f\hat{x}_{k,i}, \tag{5.4}$$

where f is a constant scalar gain, and $\hat{x}_{k,i}$ is an estimator of $x_{k,i}$ based only on observations of the i^{th} agent. More specifically, the control is a linear feedback $-f\hat{x}_{k,i}$ on the state estimate of $x_{k,i}$, while the state estimate $\hat{x}_{k,i}$ is obtained based solely on agent i's own observations $y_{k,i},\ y_{k-1,i},\ y_{k-2,i},\ y_{k-3,i},\ \cdots$.

5.4 Filtering for a Class of Certainty Equivalent Controllers

5.4.1 Previous Work

In Abedinpour Fallah *et al.* (2016), we explored the conditions under which a Luenberger like observer of the form

$$\hat{x}_{k+1,i} = (a - bf)\hat{x}_{k,i} + K(y_{k+1,i} - c(a - bf)\hat{x}_{k,i}), \tag{5.5}$$

together with a constant state estimate feedback in individual systems would be : (i) ideally optimal, (ii) at least stable. It was found that (i) was asymptotically true (as N goes to infinity) when state gain a is less than a value called a_{Nash} and which can be exactly computed. Up to a_{Nash}, the optimum control policy is the isolated (naive) Kalman gain K^* (obtained by assuming zero interference in the local measurements, i.e., setting $h = 0$ in (5.2)) combined with the Riccati dictated optimal gain f^*, with

$$K^*(a) = \frac{cP_\infty(a)}{c^2 P_\infty(a) + \sigma_v^2}, \tag{5.6}$$

where $P_\infty(a)$ is the unique positive solution of

$$c^2 P_\infty^2(a) + ((1 - a^2)\sigma_v^2 - c^2\sigma_w^2)P_\infty(a) - \sigma_w^2\sigma_v^2 = 0, \tag{5.7}$$

and

$$f^*(a, r) = \frac{ab\Sigma_\infty}{b^2\Sigma_\infty + r}, \tag{5.8}$$

where Σ_∞ is the positive solution of the algebraic Riccati equation

$$b^2\Sigma_\infty^2 + (r - a^2 r - b^2)\Sigma_\infty - r = 0. \tag{5.9}$$

There is also a_{sup} greater than a_{Nash} such that when a is between a_{Nash} and a_{sup}, one can reverse engineer a range of coefficients r for the cost functions in (5.3) for which the naive Kalman gain K^* combined with the feedback gain f^* dictated by the Riccati equation will be asymptotically optimal. Finally, past a_{sup}, no optimal control interpretation is possible any more, although there exist couples (K, f) which may still stabilize the system up to a maximum value, and only cooperatively chosen common gains can get us to approach a minimum cost. Let a_s be the limit past which constant Luenberger like observer and feedback gains can no longer stabilize the system. The current paper is a continuation of stabilization and optimality investigations for values of a past a_s.

From Abedinpour Fallah *et al.* (2016), for each fixed a it is possible to stabilize the system using a Luenberger like observer equation (5.5) and the pair (K, f) if and only if (K, f) is in a stability region denoted by $S(a)$, independent of N. Let $a_f = a - bf$, and also let $a_m(f)$ (or equivalently $a_m(a_f)$) denote the maximum value of a such that $(K^*(a), a_f) \in S(a)$ for some $a_f \in [0, 1)$. Moreover, let $a_m = \sup_{a_f \in [0,1)} a_m(a_f)$. Then we have the following numerical results.

Numerical results

The numerical results reported in this paper are obtained considering the following parameter setting: $b = c = h = 1$ and $\sigma_w = \sigma_v = \sigma_0 = 1$, with $\mathbb{E}x_{0,i} = \bar{x}_0 = 0$ for all agents $i = 1, 2, \ldots, N$. In addition, we will only deal with the case $a \geq 0$ (a symmetric property holds for the $a \leq 0$ case). The value of a and f will be specified in the different simulations.

By numerical investigation we have $a_m \approx 3.6$ which is obtained by letting a_f go to 1. Figure 5.1 is a representation of the stability regions for the assumed parameter setting when a varies from $a = 0.2$ to $a_s = 5.5$. It is observed that the stability region gradually shrinks as a increases until it all but vanishes at $a_s = 5.5$. More specifically, for all $a < a_{Nash}$, all intersections of the horizontal line $K^*(a)$ with all the vertical lines (marking values of f which, given a , satisfy (5.8) for some value of input penalty coefficient r, where r goes from 0 to infinity) belong to the stability region. For all $a \in (a_{Nash}, a_{sup})$ this holds only for large enough r while for $a = a_{sup}$ all the intersections cease to belong to the brown area. For $a = a_m$ the Kalman gain ceases to be in the stability region for all $a_f \in [0, 1)$ and, for $a = 5.5$ the stability region becomes empty.

5.4.2 Exact Optimal Bulk Filter

The goal of this section is to extend the class of candidate stabilizing control structures via *exact* optimal filtering with $h \neq 0$ in its formulation. In particular, when local state estimate feedback (5.4) is included in the i^{th} agent state equation (5.1), the result is as follows:

$$x_{k+1,i} = ax_{k,i} - bf\hat{x}_{k,i} + w_{k,i}. \tag{5.10}$$

In addition, anticipating the need to account for the influence of average states in the dynamics through the measurement equation, and letting a tilde $(\tilde{\cdot})$ indicate an averaging over

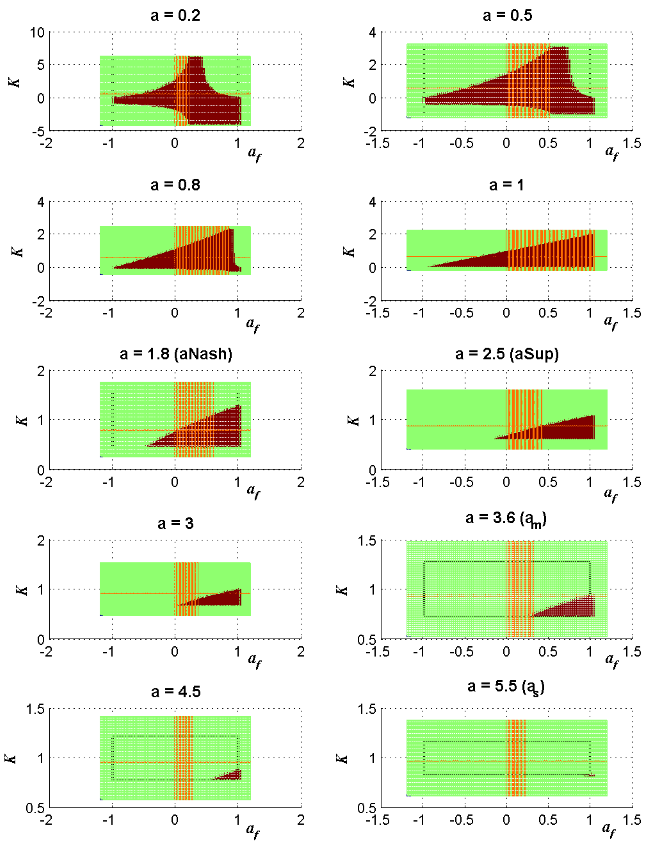

Figure 5.1 The stability regions $S(a)$ (brown shaded areas) in the (K, a_f) plane. The vertical lines represent the optimal Riccati gain $f^*(a, r)$ corresponding to all possible values of parameter r while the horizontal line is the optimal isolated (naive) Kalman filter gain $K^*(a)$.

agents operation, we define:

$$m_k = \frac{1}{N}\sum_{j=1}^{N} x_{k,j}, \quad \tilde{m}_k = \frac{1}{N}\sum_{j=1}^{N} \hat{x}_{k,j}, \tag{5.11}$$

$$\tilde{w}_k = \frac{1}{N}\sum_{j=1}^{N} w_{k,j}, \quad \tilde{v}_k = \frac{1}{N}\sum_{j=1}^{N} v_{k,j}. \tag{5.12}$$

Thus, combining (5.10), (5.11), (5.12), we obtain:

$$m_{k+1} = am_k - bf\tilde{m}_k + \tilde{w}_k \tag{5.13}$$

It appears that optimal decentralized estimation in our problem structure, does not lend itself naturally to a recursive computation. Indeed, the sufficient statistic from the past (adequate state description) appears to grow by 2 components at every step. More specifically, average agent state (averaging over all agents) and average agent state estimate enter into the dynamics of the controlled individual agents, through the $-f\hat{x}_{k,i}$ term into the dynamics (5.1), and thus the dynamics of both of these averages must be specified to complete the estimation procedure. As a result, one has to augment the dynamics of (5.1) by at least that of the agents states average term in the estimation at the first step. When optimal filtering is applied to the augmented state, computations of both the innovation term and its gain involve an expected value of the average state and the average state estimate (because the latter enters the average agent state dynamics). This yields to the apparent infinite regress effect.

A significant source of complexity in the analysis, is self dependency of filtering equations. Roughly speaking, since the averaged agent state enters into individual measurements, the effective stochasticity in a single agent's measurements depends on how precisely other agents are estimating their individual states. Thus by symmetry, the level of uncertainty in individual state estimates depends on itself. Furthermore, the straightforward recursive Kalman filter assumes that the internal and measurement sequences noises are uncorrelated with their past (white noise property). However, in the sequence of expanding state models that we need to construct for estimation purposes as the time index increases, the noise vectors are also expanding, and are partially common from one stage to another.

However, given that all noise and initial random variables are jointly Gaussian and noting that linearity is preserved in our control structure set up, optimal estimates will be linear functions of the measurements. Hence, using the classical Gaussian unbiased minimum variance estimation theory (Bagchi (1993)), we derive the exact optimal growing dimension filter whereby at every time step, all past and present available measurements are considered. In particular, let us denote by K_k the time-varying $1 \times k$ row vector of filter gains necessary for computing $\hat{x}_{k,i}$. Also let $Y_{1,i}^k$ indicate the column vector of all measurements up to time k at agent i. The minimum covariance estimator $\hat{x}_{k,i}$ minimizes the mean square estimation error

$$\Sigma_k = \mathbb{E}\left[(x_{k,i} - \hat{x}_{k,i})^2 \, |Y_{1,i}^k\right], \tag{5.14}$$

and is given by the exact growing dimension filter (the optimal Bulk filter) equations (Abedinpour Fallah *et al.* (2013a); Bagchi (1993)):

$$\hat{x}_{k,i} = \mathbb{E}\left[x_{k,i}\right] + K_k\left(Y_{1,i}^k - \mathbb{E}\left[Y_{1,i}^k\right]\right), \tag{5.15}$$

$$\Sigma_k = P_{x_{k,i}x_{k,i}} - K_k P_{x_{k,i}Y_{1,i}^k}^T, \tag{5.16}$$

with optimal time-varying gain

$$K_k = P_{x_{k,i}Y_{1,i}^k} P_{Y_{1,i}^k Y_{1,i}^k}^{-1}, \tag{5.17}$$

where

$$P_{x_{k,i}Y_{1,i}^k} = \mathbb{E}\left[(x_{k,i} - \mathbb{E}\left[x_{k,i}\right])\left(Y_{1,i}^k - \mathbb{E}\left[Y_{1,i}^k\right]\right)^T\right], \tag{5.18}$$

$$P_{Y_{1,i}^k Y_{1,i}^k} = \mathbb{E}\left[\left(Y_{1,i}^k - \mathbb{E}\left[Y_{1,i}^k\right]\right)\left(Y_{1,i}^k - \mathbb{E}\left[Y_{1,i}^k\right]\right)^T\right]. \tag{5.19}$$

The next theorem gives a semi-recursive computational scheme for the exact growing dimension filter (5.15)-(5.17), which uses all of the results from previous cycles up to time $k-1$ to compute K_k in one shot.

Theorem 5.1. *The optimal decentralized state estimator $\hat{x}_{k,i}$ is given by:*

$$\hat{x}_{k,i} = (a-bf)^k \bar{x}_0 + K_k \begin{bmatrix} y_{1,i} - (c+h)(a-bf)\bar{x}_0 \\ y_{2,i} - (c+h)(a-bf)^2\bar{x}_0 \\ \vdots \\ y_{k,i} - (c+h)(a-bf)^k\bar{x}_0 \end{bmatrix}, \tag{5.20}$$

with optimal time-varying gain K_k, obtained from Levinson-like order-updating relations given by:

$$K_k = \begin{bmatrix}(a-bf)K_{k-1} & 0\end{bmatrix} + \frac{P_{x_{k,i}y_{k,i}} - P_{x_{k,i}Y_{1,i}^{k-1}} P_{Y_{1,i}^{k-1}Y_{1,i}^{k-1}}^{-1} P_{Y_{1,i}^{k-1}y_{k,i}}}{P_{y_{k,i}y_{k,i}} - P_{Y_{1,i}^{k-1}y_{k,i}}^T P_{Y_{1,i}^{k-1}Y_{1,i}^{k-1}}^{-1} P_{Y_{1,i}^{k-1}y_{k,i}}} \begin{bmatrix} -P_{Y_{1,i}^{k-1}y_{k,i}}^T P_{Y_{1,i}^{k-1}Y_{1,i}^{k-1}}^{-1} & 1 \end{bmatrix}, \tag{5.21}$$

with $K_1 = P_{x_{1,i}y_{1,i}}/P_{y_{1,i}y_{1,i}}$*, and*

$$P^{-1}_{Y^k_{1,i}Y^k_{1,i}} = \begin{bmatrix} P^{-1}_{Y^{k-1}_{1,i}Y^{k-1}_{1,i}} & 0_{(k-1)\times 1} \\ 0_{1\times(k-1)} & 0 \end{bmatrix}$$
$$+\frac{1}{P_{y_{k,i}y_{k,i}} - P^T_{Y^{k-1}_{1,i}y_{k,i}} P^{-1}_{Y^{k-1}_{1,i}Y^{k-1}_{1,i}} P_{Y^{k-1}_{1,i}y_{k,i}}} \begin{bmatrix} -P^{-1}_{Y^{k-1}_{1,i}Y^{k-1}_{1,i}} P_{Y^{k-1}_{1,i}y_{k,i}} \\ 1 \end{bmatrix} \begin{bmatrix} -P^{-1}_{Y^{k-1}_{1,i}Y^{k-1}_{1,i}} P_{Y^{k-1}_{1,i}y_{k,i}} \\ 1 \end{bmatrix}^T, \tag{5.22}$$

where the covariances $P_{x_{k,i}y_{t,i}}$ *and* $P_{y_{k,i}y_{t,i}}$ *are respectively defined as*

$$P_{x_{k,i}y_{t,i}} = \mathbb{E}\left[(x_{k,i} - \mathbb{E}[x_{k,i}])(y_{t,i} - \mathbb{E}[y_{t,i}])^T\right], \tag{5.23}$$

$$P_{y_{k,i}y_{t,i}} = \mathbb{E}\left[(y_{k,i} - \mathbb{E}[y_{k,i}])(y_{t,i} - \mathbb{E}[y_{t,i}])^T\right], \tag{5.24}$$

for $t = 1, ..., k$*, and are obtained recursively from the following equations:*

$$P_{x_{k,i}y_{t,i}} = \begin{cases} aP_{x_{k-1,i}y_{t,i}} - bfK_{k-1} \begin{bmatrix} P_{y_{1,i}y_{t,i}} \\ \vdots \\ P_{y_{k-1,i}y_{t,i}} \end{bmatrix}, & \textit{for } t = 1, \ldots, k-1 \\ cP_{x_{k,i}x_{t,i}} + hP_{x_{k,i}m_t}, & \textit{for } t \geq k \end{cases} \tag{5.25}$$

$$P_{y_{k,i}y_{t,i}} = \begin{cases} cP_{x_{k,i}y_{t,i}} + hP_{y_{t,i}m_k}, & \textit{for } t = 1, \ldots, k-1 \\ cP_{x_{k,i}y_{k,i}} + hP_{y_{k,i}m_k} + \sigma_v^2, & \textit{for } t = k \end{cases} \tag{5.26}$$

$$P_{y_{k,i}m_t} = \begin{cases} cP_{x_{k,i}m_t} + hP_{m_k m_t}, & \textit{for } t \leq k \\ aP_{y_{k,i}m_{t-1}} - bfK_{t-1} \begin{bmatrix} (c+h)P_{y_{k,i}m_1} + P_{y_{k,i}\tilde{v}_1} \\ \vdots \\ (c+h)P_{y_{k,i}m_{t-1}} + P_{y_{k,i}\tilde{v}_{t-1}} \end{bmatrix}, & \textit{for } t > k \end{cases} \tag{5.27}$$

$$P_{x_{k,i}x_{t,i}} = \begin{cases} aP_{x_{k-1,i}x_{t,i}} - bfK_{k-1}\begin{bmatrix} P_{x_{t,i}y_{1,i}} \\ \vdots \\ P_{x_{t,i}y_{k-1,i}} \end{bmatrix},\ for\ t = 1, \ldots, k-1 \\ a^2P_{x_{k-1,i}x_{k-1,i}} - 2abfK_{k-1}\begin{bmatrix} P_{x_{k-1,i}y_{1,i}} \\ \vdots \\ P_{x_{k-1,i}y_{k-1,i}} \end{bmatrix} \\ +b^2f^2K_{k-1}P_{Y_{1,i}^{k-1}Y_{1,i}^{k-1}}K_{k-1}^T + \sigma_w^2,\ for\ t = k \end{cases} \tag{5.28}$$

$$P_{x_{k,i}m_t} = \begin{cases} aP_{x_{k-1,i}m_t} - bfK_{k-1}\begin{bmatrix} P_{y_{1,i}m_t} \\ \vdots \\ P_{y_{k-1,i}m_t} \end{bmatrix}, \quad for\ t = 1, \ldots, k-1 \\ a^2P_{x_{k-1,i}m_{t-1}} - abf(c+h)K_{k-1}\begin{bmatrix} P_{x_{k-1,i}m_1} \\ \vdots \\ P_{x_{k-1,i}m_{t-1}} \end{bmatrix} \\ -abfK_{k-1}\begin{bmatrix} P_{x_{k-1,i}\tilde{v}_1} \\ \vdots \\ P_{x_{k-1,i}\tilde{v}_{t-1}} \end{bmatrix} - abfK_{k-1}\begin{bmatrix} P_{y_{1,i}m_{t-1}} \\ \vdots \\ P_{y_{k-1,i}m_{t-1}} \end{bmatrix} \\ +b^2f^2K_{k-1}((c+h)P_{Y_{1,i}^{k-1}M_1^{t-1}} + P_{Y_{1,i}^{k-1}\tilde{V}_1^{t-1}})K_{k-1}^T \\ +aP_{x_{k-1,i}\tilde{w}_{t-1}} + aP_{m_{t-1}w_{k-1,i}} - bfK_{k-1}\begin{bmatrix} P_{y_{1,i}\tilde{w}_{t-1}} \\ \vdots \\ P_{y_{k-1,i}\tilde{w}_{t-1}} \end{bmatrix} \\ -bf(c+h)K_{k-1}\begin{bmatrix} P_{m_1w_{k-1,i}} \\ \vdots \\ P_{m_{t-1}w_{k-1,i}} \end{bmatrix} + P_{w_{k-1,i}\tilde{w}_{t-1}} \qquad for\ t \geq k \end{cases} \tag{5.29}$$

$$P_{m_k m_t} = \begin{cases} aP_{m_{k-1}m_t} - bfK_{k-1} \begin{bmatrix} (c+h)P_{m_1 m_t} + P_{m_t \tilde{v}_1} \\ \vdots \\ (c+h)P_{m_{k-1}m_t} + P_{m_t \tilde{v}_{k-1}} \end{bmatrix}, \ \mathit{for}\ t = 1, \ldots, k-1 \\ a^2 P_{m_{k-1}m_{k-1}} - 2abf(c+h)K_{k-1} \begin{bmatrix} P_{m_{k-1}m_1} \\ \vdots \\ P_{m_{k-1}m_{k-1}} \end{bmatrix} - 2abfK_{k-1} \begin{bmatrix} P_{m_{k-1}\tilde{v}_1} \\ \vdots \\ P_{m_{k-1}\tilde{v}_{k-1}} \end{bmatrix} \\ +b^2 f^2 K_{k-1}((c+h)^2 P_{M_1^{k-1} M_1^{k-1}} + (c+h)(P_{M_1^{k-1}\tilde{V}_1^{k-1}} + P^T_{M_1^{k-1}\tilde{V}_1^{k-1}}) \\ +P_{\tilde{V}_1^{k-1}\tilde{V}_1^{k-1}})K_{k-1}^T + \frac{\sigma_w^2}{N}, \quad \mathit{for}\ t = k \end{cases} \tag{5.30}$$

$$P_{y_{k,i} w_{t,i}} = cP_{x_{k,i} w_{t,i}} + hP_{m_k w_{t,i}}, \tag{5.31}$$

$$P_{m_k w_{t,i}} = \begin{cases} 0, & \mathit{for}\ t > k-1 \\ \frac{\sigma_w^2}{N}, & \mathit{for}\ t = k-1 \\ aP_{m_{k-1} w_{t,i}} - bf(c+h)K_{k-1} \begin{bmatrix} P_{m_1 w_{t,i}} \\ \vdots \\ P_{m_{k-1} w_{t,i}} \end{bmatrix}, & \mathit{for}\ t < k-1 \end{cases} \tag{5.32}$$

$$P_{x_{k,i} w_{t,i}} = \begin{cases} 0, & \mathit{for}\ t > k-1 \\ \sigma_w^2, & \mathit{for}\ t = k-1 \\ aP_{x_{k-1,i} w_{t,i}} - bfK_{k-1} \begin{bmatrix} cP_{x_{1,i} w_{t,i}} + hP_{m_1 w_{t,i}} \\ \vdots \\ cP_{x_{k-1,i} w_{t,i}} + hP_{m_{k-1} w_{t,i}} \end{bmatrix}, & \mathit{for}\ t < k-1 \end{cases} \tag{5.33}$$

$$P_{y_{k,i} \tilde{w}_t} = cP_{x_{k,i} \tilde{w}_t} + hP_{m_k \tilde{w}_t}, \tag{5.34}$$

$$P_{m_k \tilde{w}_t} = \begin{cases} 0, & \mathit{for}\ t > k-1 \\ \frac{\sigma_w^2}{N}, & \mathit{for}\ t = k-1 \\ aP_{m_{k-1} \tilde{w}_t} - bf(c+h)K_{k-1} \begin{bmatrix} P_{m_1 \tilde{w}_t} \\ \vdots \\ P_{m_{k-1} \tilde{w}_t} \end{bmatrix}, & \mathit{for}\ t < k-1 \end{cases} \tag{5.35}$$

$$P_{x_{k,i}\tilde{w}_t} = \begin{cases} 0, & \textit{for } t > k-1 \\ \frac{\sigma_w^2}{N}, & \textit{for } t = k-1 \\ aP_{x_{k-1,i}\tilde{w}_t} - bfK_{k-1} \begin{bmatrix} cP_{x_{1,i}\tilde{w}_t} + hP_{m_1\tilde{w}_t} \\ \vdots \\ cP_{x_{k-1,i}\tilde{w}_t} + hP_{m_{k-1}\tilde{w}_t} \end{bmatrix}, & \textit{for } t < k-1 \end{cases} \tag{5.36}$$

$$P_{y_{k,i}v_{t,i}} = \begin{cases} 0, & \textit{for } t > k \\ \sigma_v^2, & \textit{for } t = k \\ cP_{x_{k,i}v_{t,i}} + hP_{m_k v_{t,i}}, & \textit{for } t < k \end{cases} \tag{5.37}$$

$$P_{m_k v_{t,i}} = \begin{cases} 0, & \textit{for } t > k-1 \\ -bfK_{k-1}(k-1)\frac{\sigma_v^2}{N}, & \textit{for } t = k-1 \\ aP_{m_{k-1}v_{t,i}} - bfK_{k-1}(t)\frac{\sigma_v^2}{N} - bf(c+h)K_{k-1} \begin{bmatrix} P_{m_1 v_{t,i}} \\ \vdots \\ P_{m_{k-1}v_{t,i}} \end{bmatrix}, & \textit{for } t < k-1 \end{cases} \tag{5.38}$$

$$P_{x_{k,i}v_{t,i}} = \begin{cases} 0, & \textit{for } t > k-1 \\ -bfK_{k-1}(k-1)\sigma_v^2, & \textit{for } t = k-1 \\ aP_{x_{k-1,i}v_{t,i}} - bfK_{k-1}(t)\sigma_v^2 - bfK_{k-1} \begin{bmatrix} cP_{x_{1,i}v_{t,i}} + hP_{m_1 v_{t,i}} \\ \vdots \\ cP_{x_{k-1,i}v_{t,i}} + hP_{m_{k-1}v_{t,i}} \end{bmatrix}, & \textit{for } t < k-1 \end{cases} \tag{5.39}$$

$$P_{y_{k,i}\tilde{v}_t} = \begin{cases} 0, & \textit{for } t > k \\ \frac{\sigma_v^2}{N}, & \textit{for } t = k \\ cP_{x_{k,i}\tilde{v}_t} + hP_{m_k\tilde{v}_t}, & \textit{for } t < k \end{cases} \tag{5.40}$$

$$P_{m_k\tilde{v}_t} = \begin{cases} 0, & \mathit{for}\ t > k-1 \\ -bfK_{k-1}(k-1)\frac{\sigma_v^2}{N}, & \mathit{for}\ t = k-1 \\ aP_{m_{k-1}\tilde{v}_t} - bfK_{k-1}(t)\frac{\sigma_v^2}{N} - bf(c+h)K_{k-1}\begin{bmatrix} P_{m_1\tilde{v}_t} \\ \vdots \\ P_{m_{k-1}\tilde{v}_t} \end{bmatrix}, & \mathit{for}\ t < k-1 \end{cases} \tag{5.41}$$

$$P_{x_{k,i}\tilde{v}_t} = \begin{cases} 0, & \mathit{for}\ t > k-1 \\ -bfK_{k-1}(k-1)\frac{\sigma_v^2}{N}, & \mathit{for}\ t = k-1 \\ aP_{x_{k-1,i}\tilde{v}_t} - bfK_{k-1}(t)\frac{\sigma_v^2}{N} - bfK_{k-1}\begin{bmatrix} cP_{x_{1,i}\tilde{v}_t} + hP_{m_1\tilde{v}_t} \\ \vdots \\ cP_{x_{k-1,i}\tilde{v}_t} + hP_{m_{k-1}\tilde{v}_t} \end{bmatrix}, & \mathit{for}\ t < k-1 \end{cases} \tag{5.42}$$

where $K_k(j)$ *denotes the* j^{th} *element of* K_k*, and* M_1^k*,* $\tilde{V}_1^k$ *respectively indicate the column vector of* m *and* $\tilde{v}$ *from time* 1 *up to time* k*.*

Proof: See the Appendix.

Remark 5.1. *Note that the Bulk filter estimation at step* k *requires solving* all the intermediate steps *from* 1 *to* k*, because this is a situation of* dual control *whereby the quality of estimation at time step* k *depends on previous control actions, themselves a function of previous filtering results.*

Remark 5.2. *Note also that all the above expressions capitalize on computations already carried out at the previous time step; otherwise, the complexity of calculations would make bulk filtering estimates an essentially insurmountable task.*

The next Lemma gives the cross covariance of two arbitrary agents.

Lemma 5.1. *The cross covariance of two arbitrary agents* $x_{k,i}$ *and* $x_{k,j}$ *over time are obtained* *recursively from the following equations:*

$$P_{x_{k,i}x_{t,j}} = \begin{cases} aP_{x_{k-1,i}x_{t,j}} - bfK_{k-1}\begin{bmatrix} P_{x_{t,j}y_{1,i}} \\ \vdots \\ P_{x_{t,j}y_{k-1,i}} \end{bmatrix}, \; for\ t = 1, \ldots, k-1 \\ a^2 P_{x_{k-1,i}x_{k-1,j}} - abfK_{k-1}\begin{bmatrix} P_{x_{k-1,i}y_{1,j}} \\ \vdots \\ P_{x_{k-1,i}y_{k-1,j}} \end{bmatrix} - abfK_{k-1}\begin{bmatrix} P_{x_{k-1,j}y_{1,i}} \\ \vdots \\ P_{x_{k-1,j}y_{k-1,i}} \end{bmatrix} \\ +b^2 f^2 K_{k-1} P_{Y_{1,i}^{k-1} Y_{1,j}^{k-1}} K_{k-1}^T, \; for\ t = k \end{cases} \tag{5.43}$$

where

$$P_{x_{k,i}y_{t,j}} = \begin{cases} aP_{x_{k-1,i}y_{t,j}} - bfK_{k-1}\begin{bmatrix} P_{y_{1,i}y_{t,j}} \\ \vdots \\ P_{y_{k-1,i}y_{t,j}} \end{bmatrix}, & for\ t < k \\ cP_{x_{k,i}x_{t,j}} + hP_{x_{k,i}m_t}, & for\ t \geq k \end{cases} \tag{5.44}$$

$$P_{y_{k,i}y_{t,j}} = cP_{x_{k,i}y_{t,j}} + hP_{y_{t,j}m_k}, \qquad for\ t = 1, \ldots, k \tag{5.45}$$

Proof: See the Appendix.

5.4.3 Finite Memory Filtering Approximations

The optimal Bulk filter is an infinite impulse response (IIR) filter. Any stable IIR filter can be approximated to any desirable degree by a finite impulse response (FIR) filter (Manolakis *et al.* (2005)). In this section, we derive approximate finite-dimensional (time-varying) filters to reduce memory requirements of the Bulk filter. Let any variable with superscript (n) correspond to an approximate FIR filter of length n, where only the last n measurements are preserved. In particular, let us denote by $K_k^{(n)}$ the time-varying $1 \times n$ row vector of filter gains necessary for computing $\hat{x}_{k,i}^{(n)}$. Also let $Y_{k-n+1,i}^k$ indicate the column vector of n measurements from time $k-n+1$ up to time k at agent i, and assume zero mean for initial conditions of all agents, i.e., $\mathbb{E}x_{0,i} = \bar{x}_0 = 0, i \geq 1$. The minimum covariance estimator $\hat{x}_{k,i}^{(n)}$ minimizes the mean square estimation error

$$\Sigma_k^{(n)} = \mathbb{E}\left[\left(x_{k,i} - \hat{x}_{k,i}^{(n)}\right)^2 |Y_{k-n+1,i}^k\right], \tag{5.46}$$

and is given by (Manolakis *et al.* (2005)):

$$\hat{x}_{k,i}^{(n)} = K_k^{(n)} Y_{k-n+1,i}^k, \tag{5.47}$$

$$\Sigma_k^{(n)} = P_{x_{k,i}x_{k,i}} - K_k^{(n)} P_{x_{k,i}Y_{k-n+1,i}^k}^T, \tag{5.48}$$

with optimum time-varying gain

$$K_k^{(n)} = P_{x_{k,i}Y_{k-n+1,i}^k} P_{Y_{k-n+1,i}^k Y_{k-n+1,i}^k}^{-1}. \tag{5.49}$$

The next theorem gives a semi-recursive computational scheme for the optimum finite memory filter (5.47)-(5.49), which uses all of the results from time $k-n$ up to time $k-1$ to compute $K_k^{(n)}$ in one shot.

Theorem 5.2. *The approximate optimal decentralized state estimator $\hat{x}_{k,i}^{(n)}$ is given by:*

$$\hat{x}_{k,i}^{(n)} = K_k^{(n)} Y_{k-n+1,i}^k, \tag{5.50}$$

with optimal time-varying gain $K_k^{(n)}$, obtained from Levinson-like order-updating relations given by:

$$\begin{aligned} K_k^{(l+1)} &= \begin{bmatrix} 0 & K_k^{(l)} \end{bmatrix} \\ &+ \frac{P_{x_{k,i}y_{k-l,i}} - P_{x_{k,i}Y_{k-l+1,i}^k} P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} P_{y_{k-l,i}Y_{k-l+1,i}^k}^T}{P_{y_{k-l,i}y_{k-l,i}} - P_{y_{k-l,i}Y_{k-l+1,i}^k} P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} P_{y_{k-l,i}Y_{k-l+1,i}^k}^T} \begin{bmatrix} 1 & -P_{y_{k-l,i}Y_{k-l+1,i}^k} P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} \end{bmatrix}, \end{aligned} \tag{5.51}$$

for $l = 1, ..., n-1$, with $K_k^{(1)} = P_{x_{k,i}y_{k,i}} / P_{y_{k,i}y_{k,i}}$, and

$$\begin{aligned} P_{Y_{k-l,i}^k Y_{k-l,i}^k}^{-1} &= \begin{bmatrix} 0 & 0_{1\times l} \\ 0_{l\times 1} & P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} \end{bmatrix} + \frac{1}{P_{y_{k-l,i}y_{k-l,i}} - P_{y_{k-l,i}Y_{k-l+1,i}^k} P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} P_{y_{k-l,i}Y_{k-l+1,i}^k}^T} \\ &\begin{bmatrix} 1 \\ -P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} P_{y_{k-l,i}Y_{k-l+1,i}^k}^T \end{bmatrix} \begin{bmatrix} 1 \\ -P_{Y_{k-l+1,i}^k Y_{k-l+1,i}^k}^{-1} P_{y_{k-l,i}Y_{k-l+1,i}^k}^T \end{bmatrix}^T, \end{aligned} \tag{5.52}$$

for $l = 1, ..., n-1$, where the covariances $P_{x_{k,i}y_{t,i}}$ and $P_{y_{k,i}y_{t,i}}$ are obtained recursively from the truncated covariance expressions (5.25)-(5.42) by considering $k-n+1 \leq t \leq k$, and replacing K_k with $K_k^{(n)}$.

Proof: See the Appendix.

5.4.4 The Steady State Isolated Kalman Sequence

In this section, we obtain the equivalent sequence for the isolated Kalman filter equation in the steady-state.

Proposition 5.1. *In the stability region, the isolated Kalman filter equivalent equation in the steady-state is given by*

$$\hat{x}_{k,i} = K'_1 y_{k,i} + K'_2 y_{k-1,i} + ... + K'_k y_{1,i}, \tag{5.53}$$

where

$$K'_j = (a - bf)^{j-1}(1 - cK^*)^{j-1}K^*. \tag{5.54}$$

Proof: Let K be given as (5.6) in (5.5). By rearranging (5.5) we have:

$$\hat{x}_{k+1,i} = (a - bf)(1 - cK^*)\hat{x}_{k,i} + K^* y_{k+1,i}. \tag{5.55}$$

Then substituting

$$\hat{x}_{k,i} = K'_1 y_{k,i} + K'_2 y_{k-1,i} + ... + K'_k y_{1,i}, \tag{5.56}$$

and

$$\hat{x}_{k+1,i} = K'_1 y_{k+1,i} + K'_2 y_{k,i} + ... + K'_{k+1} y_{1,i}, \tag{5.57}$$

into (5.55) and applying the stationarity property by making the left-hand-side of the resulting equation equal to its right-hand-side, we get the fixed-point values (5.54). □

5.5 Numerical Study of Filtering and Control Performance

In this section, we numerically investigate state tracking ability of both the exact (growing dimension) bulk filter and its approximate finite dimension versions, as well as the control performance of the associated certainty equivalent controllers.

5.5.1 Stabilization Ability of Certainty Equivalent Controllers

First, we numerically demonstrate the stabilization ability of our class of certainty equivalent controllers using an arbitrary feedback gain f such that $|a_f| < 1$ (where $a_f = a - bf$) on the bulk filter estimate. In particular, Figs. 5.2, 5.3 and 5.4 respectively, show the stable behavior of a representative agent (namely, the 50^{th} agent) when $a = 10$, $a = 100$, and $a = 1000$.

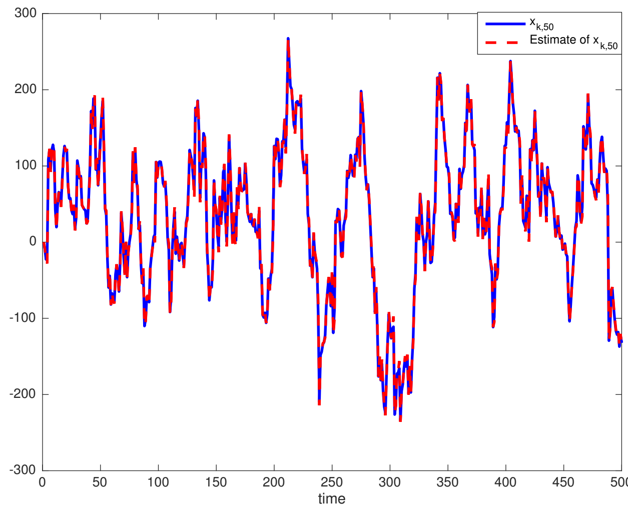

Figure 5.2 The state and its estimate when $a = 10$, $a_f = 0.9$, $\sigma_v = 1$, $N = 100$, $i = 50$.

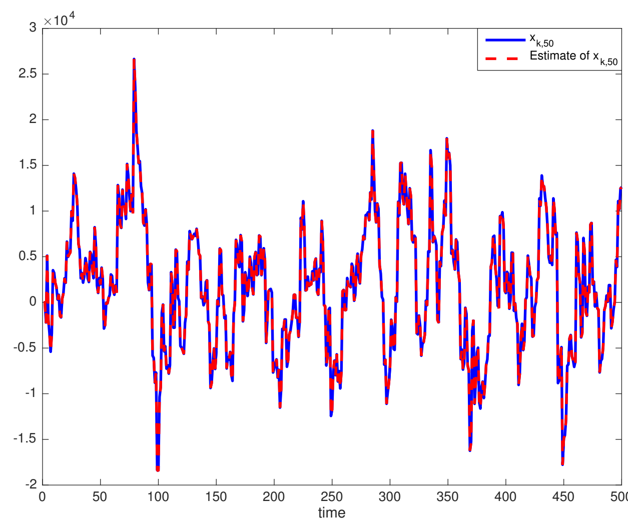

Figure 5.3 The state and its estimate when $a = 100$, $a_f = 0.9$, $\sigma_v = 1$, $N = 100$, $i = 50$.

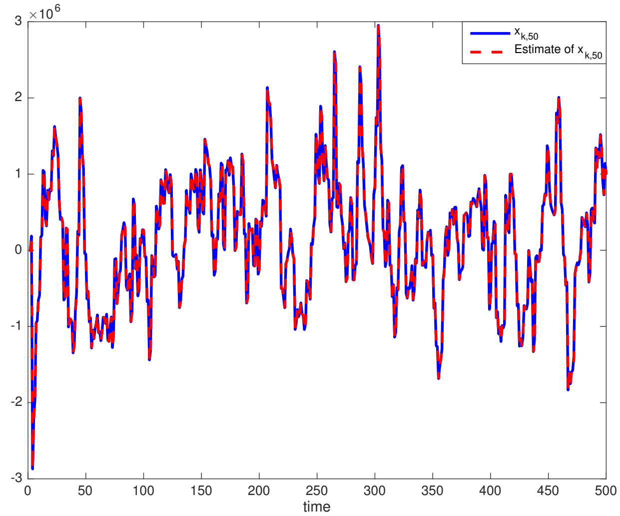

Figure 5.4 The state and its estimate when $a = 1000$, $a_f = 0.9$, $\sigma_v = 1$, $N = 100$, $i = 50$.

5.5.2 Stationarization Conditions of Optimum Filter

It is numerically observed that the optimal bulk filter gains become asymptotically stationary for $a < \bar{a}(f, N)$ or equivalently $a < \bar{a}(a_f, N)$. We shall refer to $\bar{a}(a_f, N)$ as the stationarity threshold at a_f and N. For example, when $a = 2.5$, $f = 2$, $N = 1000$, we have: $K_{Bulk} = [\ldots, -0.0006, 0.0015, 0.0018, 0.0615, 0.8628]$.
Note that the weights of the past measurements are getting smaller and smaller and disappear. The following index has been considered to investigate the stationarity of the Bulk Filter:

$$\Delta = \|K_{T-1}(1 : T-1) - K_T(2 : T)\|, \tag{5.58}$$

where $K_k(t_1 : t_2) = (K_k(t_1), K_k(t_1+1), \ldots K_k(t_2))$ comprises the elements from t_1 to t_2 of the bulk filter gain vector. In particular, the quantity Δ in (5.58) is the norm of the difference between the gains of the Bulk Filter in the last two steps of the considered time horizon $[0, T]$. If this quantity starts to increase when a is increased, it is an indication that the Bulk filter is losing its stationarity property. The values of the stationarity index Δ in (5.58) are reported in Figure 5.5 when $a_f = 0.5$ (top) and $a_f = 0.99$ (bottom) for different values of N over an horizon of $T = 1000$ steps. Moreover, the values of the thresholds $\bar{a}(a_f, N)$ and $a_m(a_f)$ for various a_f and N are reported in Tab. 5.1, where it is shown how the number of agents N affects $\bar{a}(a_f, N)$. For each fixed N, it can be observed that the threshold increases with a_f. Also, as for the dependence on N, for small values of a (namely 0.3 and 0.5) the threshold $\bar{a}(a_f, N)$ roughly decreases with N while for large values of a (namely $a = 0.8$ and $a = 0.99$) it increases. In essence, the dependence of the threshold $\bar{a}(a_f, N)$ on N remains rather weak (at least for $N \geq 100$) and the threshold is always very close to the threshold $a_m(a_f)$ where the stationary Kalman gain ceases to stabilize (5.1) for the considered a_f, under the naive filtering scheme (5.5).

Table 5.1 Thresholds $\bar{a}(a_f, N)$ and $a_m(a_f)$

a_f	$\bar{a}(a_f, N = 100)$	$\bar{a}(a_f, N = 1000)$	$\bar{a}(a_f, N = 10000)$	$a_m(a_f)$
0.3	2.48	2.44	2.38	2.36
0.5	2.79	2.8	2.76	2.74
0.8	3.25	3.35	3.34	3.304
0.99	3.55	3.72	3.72	3.67

5.5.3 Relation with the Naive Kalman Filter Approach

When the Bulk filter becomes stationary, it is interesting to investigate the relation between its steady state gain coefficients and the isolated Kalman filter equivalent equations given by (5.54). In particular, consider the following l_2 norm based discrepancy index:

$$\delta_p = \|K_{Bulk} - K_{isolated}\|, \tag{5.59}$$

where $K_{isolated} = [\dots K'_3,\ K'_2,\ K'_1]$, with $K'_j = (a - bf)^{j-1}(1 - cK^*)^{j-1}K^*$. Figure 5.6 shows the discrepancy index δ_p between the Bulk Filter gains at step T and $K_{isolated}$ as a function of a when $a_f = 0.5$ (up), 0.99 (bottom), under different values of N over an horizon of $T = 1000$ steps. It is observed that the steady state bulk filter gain with mostly a few coefficients non zero could be recovered by applying the isolated Kalman filter equivalent equations given by (5.54) provided that N is sufficiently large. For example, when $a = 2.5,\ f = 2,\ N = 10^9$, we have:

$$K_{Bulk} = [\dots \quad 0.0003 \quad 0.0039 \quad 0.0584 \quad 0.8650],$$

whereas the stationary Kalman filtering sequence as given by (5.54) is given by:

$$K_{isolated} = [\dots \quad 0.00026616 \quad 0.0039 \quad 0.0584 \quad 0.8650].$$

5.5.4 Persistently Time Varying Behavior for Large Enough a

For $a \geq \bar{a}(f, N)$ the optimal bulk filter gains remain *time-varying*. For example, when $a = 5,\ f = 4.5,\ N = 1000$ Figs. 5.7, 5.8 and 5.9 respectively, show the last entry of K_k,

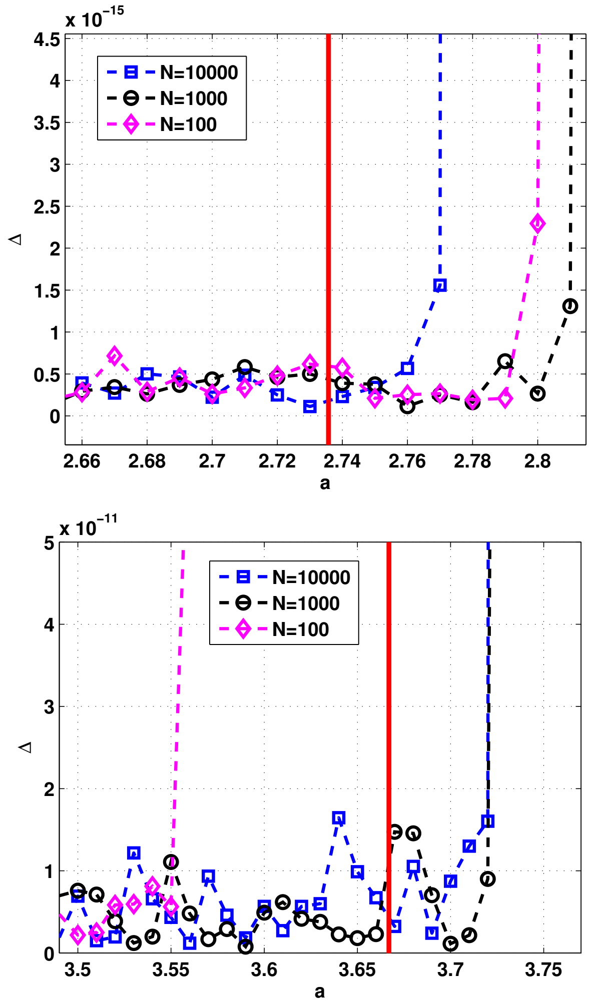

Figure 5.5 Stationarity index Δ as a function of a when $a_f = 0.5$ (top) and $a_f = 0.99$ (bottom), for different values of N over $T = 1000$ steps. The red vertical line represents the threshold $a_m(a_f)$ where $(K^*(a), a_f)$ ceases to belong to the stability region $S(a)$ in Fig. 5.1.

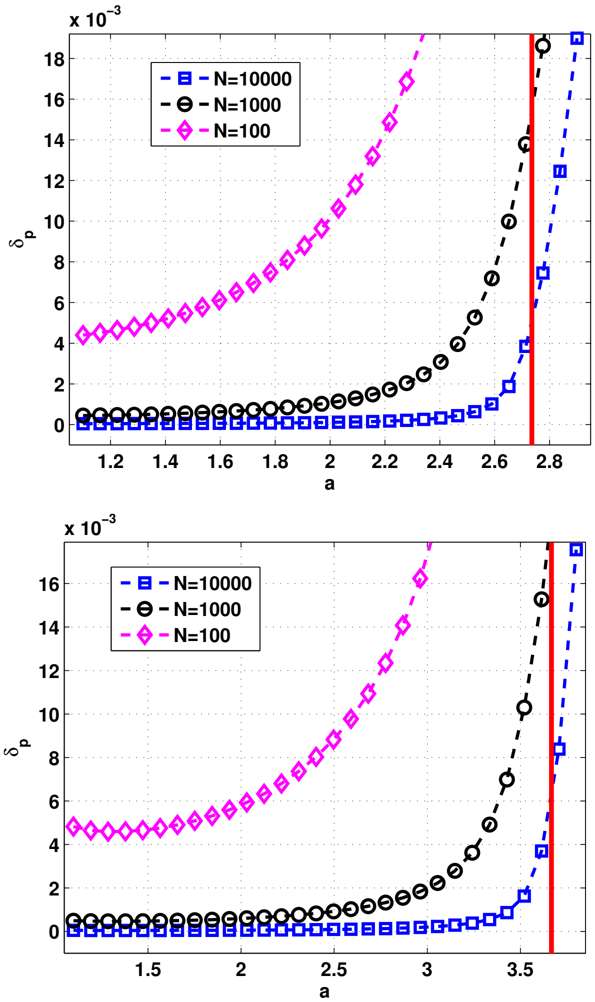

Figure 5.6 Discrepancy index δ_p between the Bulk Filter gains at step T and $K_{isolated}$ as a function of a when $a_f = 0.5$ (up), 0.99 (bottom), for different values of N over $T = 1000$ steps. The red vertical line represents the threshold $a_m(a_f)$ where $(K^*(a), a_f))$ ceases to belong to the stability region $S(a)$ in Fig. 5.1.

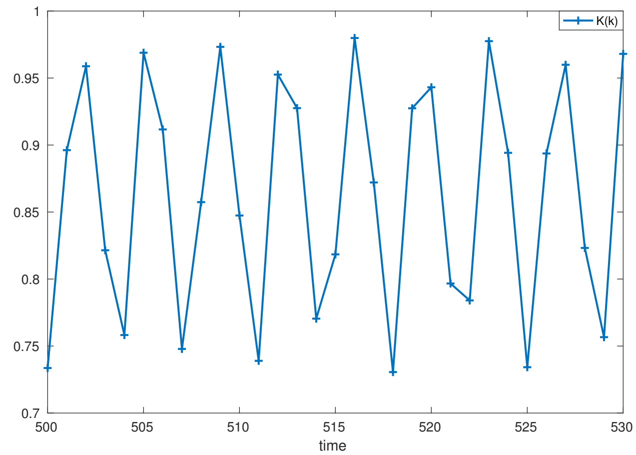

Figure 5.7 Last entry of K_k when $a = 5$, $f = 4.5$, $N = 1000$.

the cross covariance of two arbitrary agents and the mean variance. It is observed that the gains exhibit an oscillating behavior. In fact, we know that at $a = 5$ the line $k^*(5)$ does not intersect the stability region $S(a)$ in Fig. 5.1, and thus the naive Kalman filtering scheme cannot stabilize. When we apply the optimal Bulk filter, because of its accuracy and the stability of the closed loop dynamics, individual agent dynamics are stabilized and as long as the individual states remain weakly dependent, as N grows, the law of large numbers dictates that the interfering mean term in measurement equation (5.2) go to zero. At that stage, the agents are essentially independent (notice the periodic drop in interstate correlation, Fig. 5.8), and non interfering and the optimal filtering gain vector ultimately becomes that associated with the naive Kalman filter, i.e. stationary sequence (5.54), with $K^* = 0.9616$. However, we do know from Fig. 5.1 that such a filtering scheme fails to stabilize the mean dynamics, and thus after a while, the interference term grows again, thus dominating the measurement equation and creating a growing interstate cross correlation (see Fig. 5.8 again, and the coincidence of its peaks with those of the variance of m in Fig. 5.9). The interdependence of states prevents the size of N from helping in knocking out interference, and the bulk filter starts again weighing more past measurements in its estimation (this can be observed in Fig. 5.7 where the peaks of K_k roughly coincide with the minima of cross covariance terms in Fig. 5.8). Thus persistent oscillations appear in the bulk filtering gain sequences. As the degree of instability of a increases, the period of cycles gets larger and larger as the filter has to fight longer to achieve stabilization, while the trail of non negligible coefficients associated with past measurements becomes longer. At some point, a is sufficiently large that it may become difficult to clearly distinguish cycles, and filter behavior becomes quite complex to assess, although numerically, it appears to always maintain boundedness of closed loop behavior irrespective of the size of a. Also, the dependency of closed loop behavior on N becomes non monotonic, as it contributes to worsening interference when states are correlated, and quickly diminishing it when states become weakly dependent.

5.5.5 Non-Optimality of the Riccati Gain

It is observed that, in general, the Riccati gain f^* given by (5.8) is not asymptotically optimal as the number N of agents goes to infinity when using the Bulk Filter. In particular, in the range of a's for which (k^*, f) is stabilizing, the coupling term of the mean asymptotically disappears from the measurements. In that case, f^* is the optimum provided that f^* is in the range of (k^*, f) couples which still stabilize the system, and indeed the bulk filter becomes in that case equivalent to the naive Kalman filter. However, outside of the stationarization range, the quality of estimation depends on the applied feedback gain (dual effect of the control Feldbaum (1960); Witsenhausen (1968)) and this explains the non-optimality of the Riccati gain in general. For example, Figs. 5.10 and 5.11 respectively, show the average LQ cost of all the population for $a = 2$ (in stationarization region) and $a = 5$ (outside of stationarization region) using different values of N, where each point in the figures represents the average of 50 independent simulation runs and $\bigtriangledown/\bigtriangleup$ markers illustrate the standard deviation of the 50 simulations. Moreover, the vertical lines in the figures represent the Riccati gain f^*.

5.5.6 Performance of the Finite Memory Filter Approximations

Fig. 5.12 illustrates the cost defined as $\max_i |x_{T,i}|$, $1 \leq i \leq N$, i.e. the maximum value of the agent states in the last step of the simulation (saturated at 1000), obtained by adopting the finite memory Bulk filters with memory lengths of $n = 3$, $n = 5$, $n = 10$ and $n = 50$, for different values of a and a_f. It is observed that the stabilizing capability of the approximate finite memory filter is improved by increasing the memory length n.

5.6 Application to Networked Control Systems

In Abedinpour Fallah *et al.* (2016), we presented an application example for decentralized power control in code division multiple access (CDMA) cellular telephone systems with state gain $a = 1$. In this section, we present an application example for control of multi-agent systems over a CDMA network with *arbitrary* state gain a.

Following the works in Tse and Hanly (1999) and Verdú and Shamai (1999), we consider a model of a CDMA based communication and control system in the context of a large number of users with N users which share a channel and are assumed to be equally spaced on a circle

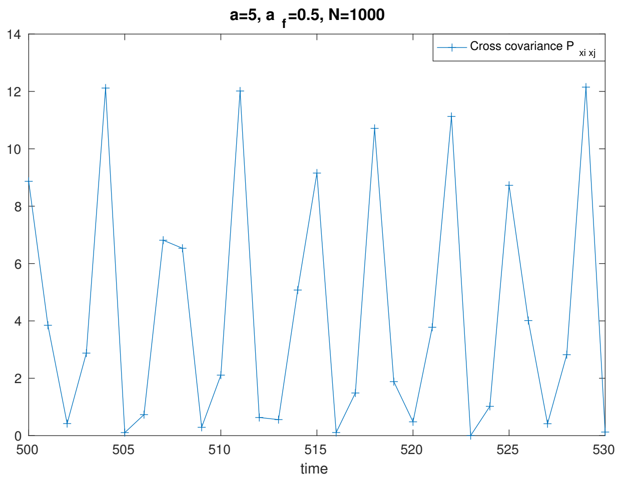

Figure 5.8 The cross covariance of two arbitrary agents when $a = 5,\ f = 4.5,\ N = 1000$.

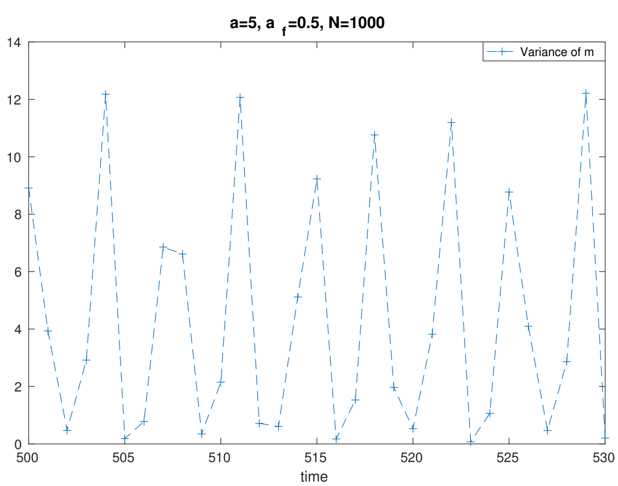

Figure 5.9 Mean variance when $a = 5,\ f = 4.5,\ N = 1000$.

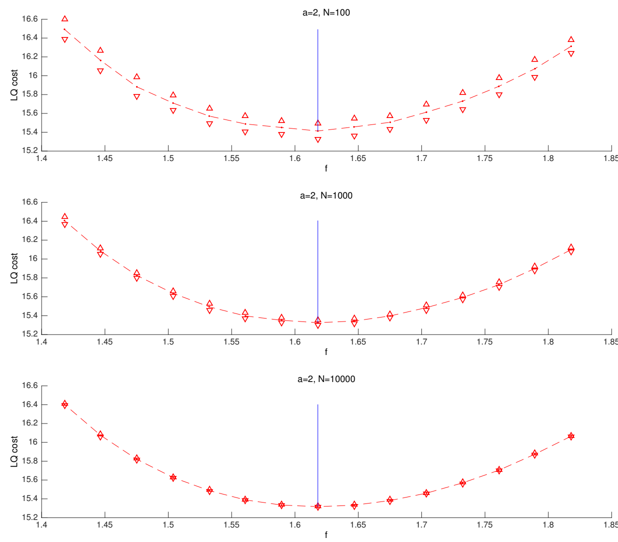

Figure 5.10 The average LQ cost of all the population for $a = 2$ and different values of N over $T = 1200$ steps, while the vertical blue line shows the Riccati gain f^{*}.

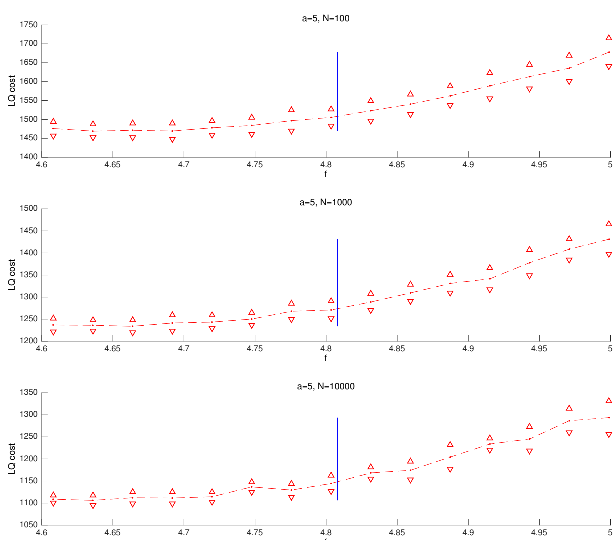

Figure 5.11 The average LQ cost of all the population for $a = 5$ and different values of N over $T = 1200$ steps, while the vertical blue line shows the Riccati gain f^{*}.

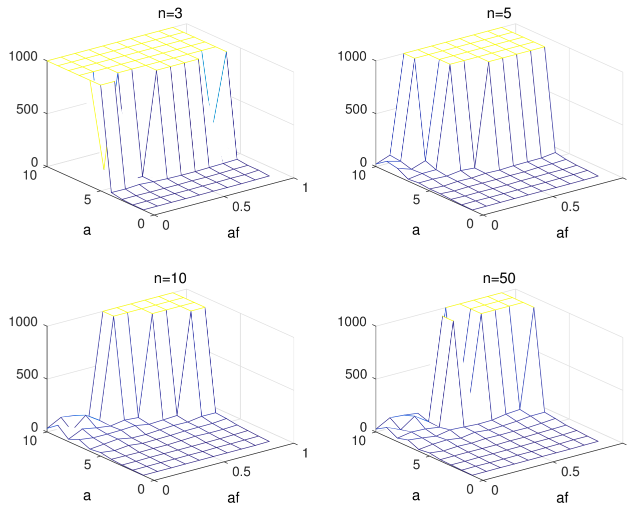

Figure 5.12 Simulative evaluation of the finite memory approximate Bulk filters with memory length of n when $N = 100$, $T = 1000$. The mesh reports the cost defined as $\max_i |x_{T,i}|$, $1 \leq i \leq N$.

around the base station, with a signal processing gain proportional to $1/N$. The base station itself sends the control signal to a collection of individual systems (users), hereon also referred to as agents. Downlink channels are considered noiseless, however the controlled individual systems are stochastic. The i^{th} mobile user of the network transmits a signal proportional to the (scalar) state $x_{k,i}$, that is to say, $\beta x_{k,i}$, where β is a constant parameter. Note that the transmitted power is proportional to $\beta^2 x_{k,i}^2$ and that the larger the state, the more energy will be involved in the transmission. The base station in turn computes the required control based on received signal which also is tainted by interference and noises. In particular, the output signal corresponding to the i^{th} user (agent) is given by:

$$y_{k,i} = \alpha\beta x_{k,i} + \frac{h'}{N}\sum_{j\neq i}^{N} \alpha\beta x_{k,j} + v_{k,i}^{th} + v'_{k,i}, \tag{5.60}$$

where $\alpha > 0$ denotes the uplink channel gain of the network, $v_{k,i}^{th}$ is the background thermal noise process and $v'_{k,i}$ is the the local observation error after transmission ($v_{k,i}^{th}$ and $v'_{k,i}$ are modeled as zero mean Gaussian random variables) [see Abedinpour Fallah *et al.* (2016); Huang *et al.* (2004); Koskie and Gajic (2006); Perreau and Anderson (2006)]. Note that the resulting signal processing gain is assumed to be h'/N. Also, the actual controlling users are assumed to be independent and simply using the base station as a communication tool. They would not want to share in any way their private information in a cooperative scheme that would allow others to identify their state.

Thus, by letting

$$c = \alpha\beta(1 - \frac{h'}{N}), \quad h = \alpha\beta h', \quad v_{k,i} = v^{th}_{k,i} + v'_{k,i}, \tag{5.61}$$

the physics of CDMA transmission viewed as a networked control system with N agents can be cast into a state space form with individual scalar dynamics described by (5.1) and partial observations given by (5.2).

5.7 Conclusion

In this paper, we have studied a class of certainty equivalent controllers with time invariant state estimate feedback gain for uniform agents that have linear stochastic individual dynamics and are coupled only through an interference term (the mean of all agent states), entering each of their individual measurement equations. The main challenge has been one of developing a decentralized filtering scheme under the considered class of feedback controllers. The optimum filters present several complicating features: (i) their form and performance is control dependent and thus a “dual" control effect is present; (ii) the filters are growing memory while no finite dimensional sufficient statistic appears within grasp. However, we have succeeded in developing a semi-recursive computational scheme which capitalizes on numerical results from all previous cycles, for otherwise numerical complexity rapidly explodes; (iii) It is impossible to produce a state estimate at some time k without proceeding sequentially, i.e., without having to compute filtering state estimates for all steps before k. We have numerically observed that the proposed estimator in combination with an arbitrary (stabilizing under perfect state observations) state estimate feedback gain, succeeds in maintaining the boundedness of the closed loop system even when individual systems are highly unstable. Moreover, we have established existence of a stationarity threshold $\bar{a}(f, N)$ past which, the optimal filter gains never stationarize, i.e. remain time-varying, and essentially periodic in the case of weakly unstable agents. An interpretation of such behavior was provided. Furthermore, we have derived approximate finite-dimensional filters to reduce memory requirements of the exact growing dimension filter.

In future work, we will attempt to mathematically establish the stabilization ability of our class of certainty equivalent controllers. Moreover, we will study the bulk filter properties as a dynamical system when no periodicity is apparent.

5.8 Appendix

Proof of Theorem 5.1. We first note that the Bulk filter (5.15)-(5.17) is initialized with $\hat{x}_{0,i} = \bar{x}_0$, and therefore we have:

$$\mathbb{E}\left[x_{k,i}\right] = (a-bf)^k \bar{x}_0, \qquad \mathbb{E}\left[m_k\right] = (a-bf)^k \bar{x}_0, \tag{5.62}$$

and

$$\begin{aligned} \mathbb{E}\left[y_{k,i}\right] &= c\mathbb{E}\left[x_{k,i}\right] + h\mathbb{E}\left[m_k\right] \\ &= (c+h)(a-bf)^k \bar{x}_0. \end{aligned} \tag{5.63}$$

Next, we partition $P_{Y_{1,i}^k Y_{1,i}^k}$ in (5.17) as

$$P_{Y_{1,i}^k Y_{1,i}^k} = \begin{bmatrix} P_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} & P_{Y_{1,i}^{k-1} y_{k,i}} \\ P^T_{Y_{1,i}^{k-1} y_{k,i}} & P_{y_{k,i} y_{k,i}} \end{bmatrix}. \tag{5.64}$$

Then applying matrix inversion by partitioning lemma Noble and Daniel (1988) we get (5.22). Moreover, by combining (5.22) and (5.17) we have:

$$\begin{aligned} K_k = &\begin{bmatrix} P_{x_{k,i} Y_{1,i}^{k-1}} P^{-1}_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} & 0 \end{bmatrix} \\ &+ \frac{-P_{x_{k,i} Y_{1,i}^{k-1}} P^{-1}_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} P_{Y_{1,i}^{k-1} y_{k,i}} + P_{x_{k,i} y_{k,i}}}{P_{y_{k,i} y_{k,i}} - P^T_{Y_{1,i}^{k-1} y_{k,i}} P^{-1}_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} P_{Y_{1,i}^{k-1} y_{k,i}}} \begin{bmatrix} -P^T_{Y_{1,i}^{k-1} y_{k,i}} P^{-1}_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} & 1 \end{bmatrix}. \end{aligned} \tag{5.65}$$

Furthermore, we note that for $t = 1, ..., k-1$

$$\begin{aligned} P_{x_{k,i} y_{t,i}} &= \mathbb{E}[(a x_{k-1,i} - bf \hat{x}_{k-1,i} + w_{k-1,i} - \mathbb{E}[x_{k,i}])(y_{t,i} - \mathbb{E}[y_{t,i}])] \\ &= a P_{x_{k-1,i} y_{t,i}} - bf K_{k-1} P_{Y_{1,i}^{k-1} y_{t,i}}, \end{aligned} \tag{5.66}$$

therefore,

$$\begin{aligned} P_{x_{k,i} Y_{1,i}^{k-1}} P^{-1}_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} &= (a P_{x_{k-1,i} Y_{1,i}^{k-1}} - bf K_{k-1} P_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}}) P^{-1}_{Y_{1,i}^{k-1} Y_{1,i}^{k-1}} \\ &= (a-bf) K_{k-1}. \end{aligned} \tag{5.67}$$

Thus, combining (5.65) and (5.67) yields (5.21).

Next, we show the derivations of the covariance expressions (5.25)-(5.42). In particular, for

$t = 1, \ldots, k-1$ we have:

$$\begin{aligned} P_{x_{k,i}y_{t,i}} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(y_{t,i} - \mathbb{E}[y_{t,i}])] \\ =& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})(y_{t,i} - \mathbb{E}[y_{t,i}])] \end{aligned} \tag{5.68}$$

and for $t \geq k$ we have:

$$\begin{aligned} P_{x_{k,i}y_{t,i}} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(y_{t,i} - \mathbb{E}[y_{t,i}])] \\ =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(c(x_{t,i} - \mathbb{E}[x_{t,i}]) + h(m_t - \mathbb{E}[m_t]) + v_{t,i})] \end{aligned} \tag{5.69}$$

thus (5.68) and (5.69) yield (5.25). Also,

$$\begin{aligned} P_{y_{k,i}y_{t,i}} =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])(y_{t,i} - \mathbb{E}[y_{t,i}])] \\ =& \mathbb{E}[(c(x_{k,i} - \mathbb{E}[x_{k,i}]) + h(m_k - \mathbb{E}[m_k]) + v_{k,i})(y_{t,i} - \mathbb{E}[y_{t,i}])] \end{aligned} \tag{5.70}$$

which yields (5.26). Moreover, for $t \leq k$ we have:

$$\begin{aligned} P_{y_{k,i}m_t} =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])(m_t - \mathbb{E}[m_t])] \\ =& \mathbb{E}[(c(x_{k,i} - \mathbb{E}[x_{k,i}]) + h(m_k - \mathbb{E}[m_k]) + v_{k,i})(m_t - \mathbb{E}[m_t])] \end{aligned} \tag{5.71}$$

and for $t > k$ we have:

$$\begin{aligned} P_{y_{k,i}m_t} =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])(m_t - \mathbb{E}[m_t])] \\ =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])(a(m_{t-1} - \mathbb{E}[m_{t-1}]) - bfK_{t-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{t-1} - \mathbb{E}[m_{t-1}]) + \tilde{v}_{t-1} \end{bmatrix} \\ & + \tilde{w}_{t-1})], \end{aligned} \tag{5.72}$$

thus (5.71) and (5.72) yield (5.27). For $t = 1, \ldots, k-1$ we have:

$$\begin{aligned} P_{x_{k,i}x_{t,i}} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(x_{t,i} - \mathbb{E}[x_{t,i}])] \\ =& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})(x_{t,i} - \mathbb{E}[x_{t,i}])] \end{aligned} \tag{5.73}$$

and for $t = k$,

$$P_{x_{k,i}x_{t,i}} = \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})^2] \tag{5.74}$$

thus (5.73) and (5.74) yield (5.28). For $t = 1, \ldots, k-1$ we have:

$$\begin{aligned} P_{x_{k,i}m_t} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(m_t - \mathbb{E}[m_t])] \\ =& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})(m_t - \mathbb{E}[m_t])] \end{aligned} \tag{5.75}$$

and for $t \geq k$,

$$\begin{aligned} P_{x_{k,i}m_t} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(m_t - \mathbb{E}[m_t])] \\ =& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i}) \\ & (a(m_{t-1} - \mathbb{E}[m_{t-1}]) - bfK_{k-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{t-1} - \mathbb{E}[m_{t-1}]) + \tilde{v}_{t-1} \end{bmatrix} + \tilde{w}_{t-1})] \end{aligned} \tag{5.76}$$

thus (5.75) and (5.76) yield (5.29). For $t = 1, \ldots, k-1$ we have:

$$\begin{aligned}
P_{m_k m_t} =& \mathbb{E}[(m_k - \mathbb{E}[m_k])(m_t - \mathbb{E}[m_t])] \\
=& \mathbb{E}[(a(m_{k-1} - \mathbb{E}[m_{k-1}]) - bfK_{k-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{k-1} - \mathbb{E}[m_{k-1}]) + \tilde{v}_{k-1} \end{bmatrix} + \tilde{w}_{k-1}) \\
& (m_t - \mathbb{E}[m_t])]
\end{aligned} \tag{5.77}$$

and for $t = k$,

$$\begin{aligned}
P_{m_k m_t} =& \mathbb{E}[(m_k - \mathbb{E}[m_k])^2] \\
=& \mathbb{E}[(a(m_{k-1} - \mathbb{E}[m_{k-1}]) - bfK_{k-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{k-1} - \mathbb{E}[m_{k-1}]) + \tilde{v}_{k-1} \end{bmatrix} + \tilde{w}_{k-1})^2]
\end{aligned} \tag{5.78}$$

thus (5.77) and (5.78) yield (5.30). Furthermore,

$$\begin{aligned}
P_{y_{k,i} w_{t,i}} =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])w_{t,i}] \\
=& \mathbb{E}[(c(x_{k,i} - \mathbb{E}[x_{k,i}]) + h(m_k - \mathbb{E}[m_k]) + v_{k,i})w_{t,i}]
\end{aligned} \tag{5.79}$$

which gives (5.31).

$$\begin{aligned}
P_{m_k w_{t,i}} =& \mathbb{E}[(m_k - \mathbb{E}[m_k])w_{t,i}] \\
=& \mathbb{E}[(a(m_{k-1} - \mathbb{E}[m_{k-1}]) - bfK_{k-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{k-1} - \mathbb{E}[m_{k-1}]) + \tilde{v}_{k-1} \end{bmatrix} + \tilde{w}_{k-1})w_{t,i}] \\
=& \mathbb{E}[aP_{m_{k-1} w_{t,i}} - bf(c+h)K_{k-1} \begin{bmatrix} P_{m_1 w_{t,i}} \\ \vdots \\ P_{m_{k-1} w_{t,i}} \end{bmatrix} + P_{\tilde{w}_{k-1} w_{t,i}}]
\end{aligned} \tag{5.80}$$

which gives (5.32).

$$\begin{aligned}P_{x_{k,i}w_{t,i}} =&\mathbb{E}[(x_{k,i}-\mathbb{E}[x_{k,i}])w_{t,i}]\\ =&\mathbb{E}[(a(x_{k-1,i}-\mathbb{E}[x_{k-1,i}])-bfK_{k-1}\begin{bmatrix}y_{1,i}-\mathbb{E}[y_{1,i}]\\ \vdots\\ y_{k-1,i}-\mathbb{E}[y_{k-1,i}]\end{bmatrix}+w_{k-1,i})w_{t,i}]\\ =&\mathbb{E}[aP_{x_{k-1,i}w_{t,i}}-bfK_{k-1}\begin{bmatrix}P_{y_{1,i}w_{t,i}}\\ \vdots\\ P_{y_{k-1,i}w_{t,i}}\end{bmatrix}+P_{w_{k-1,i}w_{t,i}}]\end{aligned} \tag{5.81}$$

which gives (5.33).

$$\begin{aligned}P_{y_{k,i}\tilde{w}_t} =&\mathbb{E}[(y_{k,i}-\mathbb{E}[y_{k,i}])\tilde{w}_t]\\ =&\mathbb{E}[(c(x_{k,i}-\mathbb{E}[x_{k,i}])+h(m_k-\mathbb{E}[m_k])+v_{k,i})\tilde{w}_t]\end{aligned} \tag{5.82}$$

which gives (5.34).

$$\begin{aligned}P_{m_k\tilde{w}_t} =&\mathbb{E}[(m_k-\mathbb{E}[m_k])\tilde{w}_t]\\ =&\mathbb{E}[(a(m_{k-1}-\mathbb{E}[m_{k-1}])-bfK_{k-1}\begin{bmatrix}(c+h)(m_1-\mathbb{E}[m_1])+\tilde{v}_1\\ \vdots\\ (c+h)(m_{k-1}-\mathbb{E}[m_{k-1}])+\tilde{v}_{k-1}\end{bmatrix}+\tilde{w}_{k-1})\tilde{w}_t]\\ =&\mathbb{E}[aP_{m_{k-1}\tilde{w}_t}-bf(c+h)K_{k-1}\begin{bmatrix}P_{m_1\tilde{w}_t}\\ \vdots\\ P_{m_{k-1}\tilde{w}_t}\end{bmatrix}+P_{\tilde{w}_{k-1}\tilde{w}_t}]\end{aligned} \tag{5.83}$$

which gives (5.35).

$$\begin{aligned}P_{x_{k,i}\tilde{w}_t} =&\mathbb{E}[(x_{k,i}-\mathbb{E}[x_{k,i}])\tilde{w}_t]\\ =&\mathbb{E}[(a(x_{k-1,i}-\mathbb{E}[x_{k-1,i}])-bfK_{k-1}\begin{bmatrix}y_{1,i}-\mathbb{E}[y_{1,i}]\\ \vdots\\ y_{k-1,i}-\mathbb{E}[y_{k-1,i}]\end{bmatrix}+w_{k-1,i})\tilde{w}_t]\\ =&\mathbb{E}[aP_{x_{k-1,i}\tilde{w}_t}-bfK_{k-1}\begin{bmatrix}P_{y_{1,i}\tilde{w}_t}\\ \vdots\\ P_{y_{k-1,i}\tilde{w}_t}\end{bmatrix}+P_{w_{k-1,i}\tilde{w}_t}]\end{aligned} \tag{5.84}$$

which gives (5.36).

$$\begin{aligned} P_{y_{k,i}v_{t,i}} =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])v_{t,i}] \\ =& \mathbb{E}[(c(x_{k,i} - \mathbb{E}[x_{k,i}]) + h(m_k - \mathbb{E}[m_k]) + v_{k,i})v_{t,i}] \end{aligned} \tag{5.85}$$

which gives (5.37).

$$\begin{aligned} P_{m_k v_{t,i}} =& \mathbb{E}[(m_k - \mathbb{E}[m_k])v_{t,i}] \\ =& \mathbb{E}[(a(m_{k-1} - \mathbb{E}[m_{k-1}]) - bfK_{k-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{k-1} - \mathbb{E}[m_{k-1}]) + \tilde{v}_{k-1} \end{bmatrix} + \tilde{w}_{k-1})v_{t,i}] \\ =& aP_{m_{k-1}v_{t,i}} - bfK_{k-1} \begin{bmatrix} (c+h)P_{m_1 v_{t,i}} + P_{\tilde{v}_1 v_{t,i}} \\ \vdots \\ (c+h)P_{m_{k-1}v_{t,i}} + P_{\tilde{v}_{k-1}v_{t,i}} \end{bmatrix}, \end{aligned} \tag{5.86}$$

which gives (5.38).

$$\begin{aligned} P_{x_{k,i}v_{t,i}} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])v_{t,i}] \\ =& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})v_{t,i}] \\ =& aP_{x_{k-1,i}v_{t,i}} - bfK_{k-1} \begin{bmatrix} P_{y_{1,i}v_{t,i}} \\ \vdots \\ P_{y_{k-1,i}v_{t,i}} \end{bmatrix} \end{aligned} \tag{5.87}$$

which gives (5.39).

$$\begin{aligned} P_{y_{k,i}\tilde{v}_t} =& \mathbb{E}[(y_{k,i} - \mathbb{E}[y_{k,i}])\tilde{v}_t] \\ =& \mathbb{E}[(c(x_{k,i} - \mathbb{E}[x_{k,i}]) + h(m_k - \mathbb{E}[m_k]) + v_{k,i})\tilde{v}_t], \end{aligned} \tag{5.88}$$

which gives (5.40).

$$
\begin{aligned}
P_{m_k\tilde{v}_t} =& \mathbb{E}[(m_k - \mathbb{E}[m_k])\tilde{v}_t] \\
=& \mathbb{E}[(a(m_{k-1} - \mathbb{E}[m_{k-1}]) - bfK_{k-1} \begin{bmatrix} (c+h)(m_1 - \mathbb{E}[m_1]) + \tilde{v}_1 \\ \vdots \\ (c+h)(m_{k-1} - \mathbb{E}[m_{k-1}]) + \tilde{v}_{k-1} \end{bmatrix} + \tilde{w}_{k-1})\tilde{v}_t] \\
=& aP_{m_{k-1}\tilde{v}_t} - bf(c+h)K_{k-1} \begin{bmatrix} P_{m_1\tilde{v}_t} \\ \vdots \\ P_{m_{k-1}\tilde{v}_t} \end{bmatrix} + P_{\tilde{w}_{k-1}\tilde{v}_t},
\end{aligned} \tag{5.89}
$$

which gives (5.41).

$$
\begin{aligned}
P_{x_{k,i}\tilde{v}_t} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])\tilde{v}_t] \\
=& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})\tilde{v}_t] \\
=& aP_{x_{k-1,i}\tilde{v}_t} - bfK_{k-1} \begin{bmatrix} P_{y_{1,i}\tilde{v}_t} \\ \vdots \\ P_{y_{k-1,i}\tilde{v}_t} \end{bmatrix} + P_{w_{k-1,i}\tilde{v}_t}
\end{aligned} \tag{5.90}
$$

which gives (5.42). This concludes the proof. □

Proof of Lemma 5.1. For $t = 1, \dots, k-1$ we have:

$$
\begin{aligned}
P_{x_{k,i}x_{t,j}} =& \mathbb{E}[(x_{k,i} - \mathbb{E}[x_{k,i}])(x_{t,j} - \mathbb{E}[x_{t,j}])] \\
=& \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})(x_{t,j} - \mathbb{E}[x_{t,j}])]
\end{aligned} \tag{5.91}
$$

and for $t = k$, we have:

$$P_{x_{k,i}x_{k,j}} = \mathbb{E}[(a(x_{k-1,i} - \mathbb{E}[x_{k-1,i}]) - bfK_{k-1} \begin{bmatrix} y_{1,i} - \mathbb{E}[y_{1,i}] \\ \vdots \\ y_{k-1,i} - \mathbb{E}[y_{k-1,i}] \end{bmatrix} + w_{k-1,i})$$
$$(a(x_{k-1,j} - \mathbb{E}[x_{k-1,j}]) - bfK_{k-1} \begin{bmatrix} y_{1,j} - \mathbb{E}[y_{1,j}] \\ \vdots \\ y_{k-1,j} - \mathbb{E}[y_{k-1,j}] \end{bmatrix} + w_{k-1,j})] \tag{5.92}$$

thus (5.91) and (5.92) yield (5.43). The rest of the proof is similar to that of Theorem 5.1. $\square$

Proof of Theorem 5.2. First, we partition $P_{Y^k_{k-l,i}Y^k_{k-l,i}}$ in (5.49) as follows:

$$P_{Y^k_{k-l,i}Y^k_{k-l,i}} = \begin{bmatrix} P_{y_{k-l,i}y_{k-l,i}} & P_{y_{k-l,i}Y^k_{k-l+1,i}} \\ P^T_{y_{k-l,i}Y^k_{k-l+1,i}} & P_{Y^k_{k-l+1,i}Y^k_{k-l+1,i}} \end{bmatrix}. \tag{5.93}$$

Then applying matrix inversion by partitioning lemma Noble and Daniel (1988) we get (5.52). Moreover, combining (5.49) and (5.52) yields (5.51). The rest of the proof is similar to that of Theorem 5.1. $\square$

CHAPTER 6 ARTICLE 3 : INTERFERENCE INDUCED GAMES IN NETWORKED CONTROL SYSTEMS AND A CLASS OF DUAL CONTROL SOLUTIONS

Mehdi Abedinpour Fallah, Roland P. Malhamé and Francesco Martinelli
Accepted for presentation at *the 20th IFAC World Congress*, Toulouse, France, July 2017

6.1 Abstract

Networked control systems must use communication links between control hubs and distributed components, possibly both to observe component states, and to send control commands. We consider a model of a CDMA based communication and control system where the signal sent from the components to the base station, acting as the control hub, is proportional to their (scalar) state, and in turn, the base station sends back the required control commands to the components. The systems are linear, and commands are constrained to be linear, possibly time varying feedback laws on current and a limited set of recent measurements. However, the individual measurements as decoded by the base station include interference terms from the set of all other components, and this inadvertently creates an interference induced game situation. The consequence is that controls have dual effects: they steer individual systems, but they can also help create additional interference. We propose an algorithm which accounts for a combination of control and estimation costs to compute symmetric Nash equilibria if they exist.

6.2 Introduction

Networked Control System (NCS) refers to a decentralized control system in which the components are connected through real-time communication channels or a data network. Thus, there may be a data link between the sensors (which collect information), the controllers (which make decisions), and the actuators (which execute the controller commands); and the sensors, the controllers, and the plant themselves could be geographically separated (Yüksel and Başar (2013)). Game theory has emerged as a well-established discipline capable of providing a resourceful and effective framework for addressing control of large scale and distributed networked systems.

Following the works in Tse and Hanly (1999) and Verdú and Shamai (1999), we consider a model of a CDMA based communication and control system in the context of a large number of users with N users which share a channel and are assumed to be equally spaced on a circle around the base station, with a signal processing gain proportional to $1/N$. The base station itself sends the control signal to a collection of individual systems (users), hereon also referred to as agents. Downlink channels are considered noiseless, however the controlled individual systems are stochastic. The i^{th} mobile user of the network transmits a signal proportional to the (scalar) state $x_{k,i}$, that is to say, $\beta x_{k,i}$, where β is a constant parameter. Note that the transmitted power is proportional to $\beta^2 x_{k,i}^2$ and that the larger the state, the more energy will be involved in the transmission. The base station in turn computes the required control based on received signal which also is tainted by interference and noises. In particular, the output signal corresponding to the i^{th} user (agent) is given by:

$$y_{k,i} = \alpha\beta x_{k,i} + \frac{h'}{N}\sum_{j\neq i}^{N} \alpha\beta x_{k,j} + v_{k,i}^{th} + v'_{k,i}, \tag{6.1}$$

where $\alpha > 0$ denotes the uplink channel gain of the network, $v_{k,i}^{th}$ is the background thermal noise process and $v'_{k,i}$ is the the local observation error after transmission ($v_{k,i}^{th}$ and $v'_{k,i}$ are modeled as zero mean Gaussian random variables) [see Abedinpour Fallah *et al.* (2016); Huang *et al.* (2004); Koskie and Gajic (2006); Perreau and Anderson (2006)]. Note that the resulting signal processing gain is assumed to be h'/N. Also, the actual controlling users are assumed to be independent and simply using the base station as a communication tool. They would not want to share in any way their private information in a cooperative scheme that would allow others to identify their state.

Recently there have been research efforts to treat the power control problem from base station to individual cellular phone through a game theoretic view. In particular, a non linear model of the channel is used by Aziz and Caines (2017) and they formulate a mean field game problem to find Nash equilibrium strategies.

In this paper, we wish to use CDMA technology to achieve distributed control in a particular way over a network. It turns out that interference due to the convergence of information signals to the base station creates a game situation, in which control laws both steer the system and affect the quality of observations, thus creating a dual control environment (for example, see Feldbaum (1960); Kim and Rock (2006)). We propose an approach to characterize potential Nash equilibrium decentralized policies within a restricted class, that of linear time varying output feedback policies involving a limited record of the most recent measurements, despite this complex dual control environment.

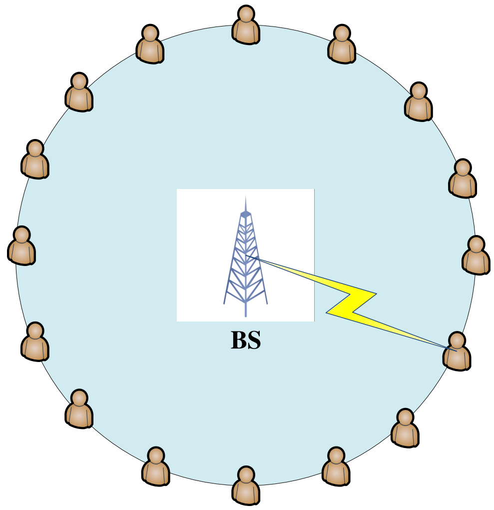

Figure 6.1 N users using CDMA technology

The rest of the paper is organized as follows. The problem is defined and formulated in state space in Section 6.3. Section 6.4 presents the main results concerning computation of symmetric Nash equilibrium using output feedback dynamic programming and Kalman filter based estimation. Section 6.5 illustrates some numerical simulation results. Concluding remarks are stated in Section 6.6.

6.3 State space model formulation and problem statement

By letting

$$c = \alpha\beta(1 - \frac{h'}{N}), \quad h = \alpha\beta h', \quad v_{k,i} = v_{k,i}^{th} + v'_{k,i}, \tag{6.2}$$

the physics of CDMA transmission viewed as a networked control system with N agents can be cast into a state space form with individual scalar dynamics described by

$$x_{k+1,i} = ax_{k,i} + bu_{k,i} + w_{k,i} \tag{6.3}$$

and partial scalar state observations given by:

$$y_{k,i} = cx_{k,i} + h\left(\frac{1}{N}\sum_{j=1}^{N} x_{k,j}\right) + v_{k,i} \tag{6.4}$$

for $k \geq 0$ and $1 \leq i \leq N$, where $x_{k,i}, u_{k,i}, y_{k,i} \in \mathbb{R}$ are the state, the control input and the measured output of the i^{th} agent, respectively. The random variables $w_{k,i} \sim \mathcal{N}(0, \sigma_w^2)$ and $v_{k,i} \sim \mathcal{N}(0, \sigma_v^2)$ represent independent Gaussian white noises at different times k and at different agents i. The Gaussian initial conditions $x_{0,i} \sim \mathcal{N}(\bar{x}_0, \sigma_0^2)$ are mutually independent and are also independent of $\{w_{k,i}, v_{k,i}, 1 \le i \le N,\ k \ge 0\}$. σ_w^2, σ_v^2 and σ_0^2 denote the variance of $w_{k,i}$, $v_{k,i}$ and $x_{0,i}$, respectively. Moreover, a is a scalar parameter and $b, c, h > 0$ are positive scalar parameters. Furthermore, the individual cost function for each agent is given by

$$J_i \triangleq \mathbb{E}\left[\sum_{k=0}^{T-1}(x_{k,i}^2 + ru_{k,i}^2) + x_{T,i}^2\right], \tag{6.5}$$

where $r > 0$ is a positive scalar parameter.

In Abedinpour Fallah *et al.* (2013a, 2014), and for the infinite horizon problem, the class of so-called separated policies (i.e. involving a constant gain feedback on the optimal local state estimate) has been considered; it has been shown that if individual agent dynamics are “sufficiently stable”, this class includes a particular feedback gain corresponding to an asymptotic in population size Nash equilibrium. Furthermore, it has been numerically observed that the class of separated policies appears to always contain stabilizing gain ranges *irrespective of the degree of instability of individual dynamics*, while predictably so, the corresponding optimal filters are linear functions of the current and past observations, however with coefficients which remain time varying. The latter observation leads us to consider existence of potential Nash equilibria, however within the special class of growing dimension time varying linear output feedback control policies given by

$$u_{k,i} = -f_{1,k}y_{k,i} - f_{2,k}y_{k-1,i} - \ldots - f_{k-1,k}y_{2,i} - f_{k,k}y_{1,i}, \tag{6.6}$$

where $f_{j,k}$ are time varying scalar gains. In this paper, we propose a methodology for the computation of such policies for a finite horizon problem (which obviously we could make as large as we wish). For analytical tractability and computability, we shall further narrow the considered class of candidate policies to that of linear output time varying feedback policies, with a limited look back at the most recent two measurements. One can make the class wider by picking more measurements, and having a growing vector of time varying gains as measurements accumulate. Thus, in the rest of the paper, we consider the problem of synthesizing Nash equilibrium strategies within a class of time varying linear output feedback control laws parameterized as follows:

$$u_{k,i} = -f_{1,k}y_{k,i} - f_{2,k}y_{k-1,i}, \tag{6.7}$$

i.e. the dependence is restricted to the most recent two measurements.

6.4 Computation of Symmetric Nash Equilibrium

The idea of the calculation is based on a “generic" agent i_0 attempting to develop a consistency principle for identifying the likely Nash candidate policies within the restricted class of interest, through the following steps:

— On a finite time horizon T, fix for every k the sequence of output feedback gains $(f_{1,k}, f_{2,k})$ to be considered as a candidate.

— Assume everyone except agent i_0 uses that feedback strategy.

— Consider N large enough that the actions of agent i_0 have negligible impact on the policies of other agents (the standard starting point of mean field games analysis - see [Huang *et al.* (2007)] for example , i.e. decoupling of the mass from the individual). Indeed, while an agent's input does not affect the quality of its estimates in the linear case, in reality the feedback gains used by i_0 affect its state and through interference will affect the ability of other agents to estimate their own state. This in turn impacts the interference perceived by i_0

— Consider the optimal control problem to be solved by individual agent i_0. Note that when solving for its best response to other agents' actions using a dynamic programming principle working backwards in time at some time k, based on the results at time $(k+1)$, it is assumed in light of the previous discussion, and for estimation purposes that up to time k the agent has been using exactly the same feedback policy as the other agents.

— Under such conditions, agent i_0 calculates the optimal feedback gains at time k, and of course, a necessary condition for the posited feedback sequence to be Nash is that the agent in question recovers the feedback gains he assumed optimal in the first place, i.e. *a fixed point result.* In this process, unlike the standard LQG dynamic programming solution, *the estimation error cost depends on the chosen feedback gains*, and thus also the structure of the fixed point equations one needs to satisfy. Also, note that in positing the quadratic structure of the optimal cost to go, we do neglect any dependencies on measurements beyond the latest two (The most general analysis would instead need to deal with optimal cost to go quadratic dependencies on a growing dimension vector of measurements).

We proceed to study the dynamical model of agent i_0. In particular, define:

$$m_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} x_{k,j}, \quad k = 0, 1, ..., T, \tag{6.8}$$

$$\tilde{w}_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} w_{k,j}, \quad \tilde{v}_k^- = \frac{1}{N}\sum_{j\neq i_0}^{N} v_{k,j}, \tag{6.9}$$

and assume that all other agents (except agent i_0) use (6.7). Thus, including (6.7) in (6.3) and combining (6.3), (6.4), (6.8), (6.9), the 5^{th} order model of the closed-loop dynamics can be expressed as:

$$X_{k+1,i_0} = A_{d,k}X_{k,i_0} + B_d u_{k,i_0} + D_{d,k}W_{k,i_0}, \tag{6.10}$$

with

$$y_{k,i_0} = H_d X_{k,i_0} + v_{k,i_0}, \tag{6.11}$$

where the augmented state is

$$X_{k,i_0} = [x_{k,i_0}, x_{k-1,i_0}, m_k^-, m_{k-1}^-, \tilde{v}_{k-1}^-]^T, \tag{6.12}$$

and matrix $A_{d,k}$ is given by

$$A_{d,k} = \begin{bmatrix} a & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ -bf_{1,k}\frac{(N-1)}{N^2}h & -bf_{2,k}\frac{(N-1)}{N^2}h & a'_{3,3} & a'_{3,4} & -bf_{2,k} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{6.13}$$

with

$$a'_{3,3} = a - bf_{1,k}(c + \frac{(N-1)}{N}h), \tag{6.14}$$

$$a'_{3,4} = -bf_{2,k}(c + \frac{(N-1)}{N}h), \tag{6.15}$$

and also we have:

$$B_d = \begin{bmatrix} b \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad D_{d,k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -bf_{1,k} \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad H_d = \begin{bmatrix} c + \frac{h}{N} \\ 0 \\ h \\ 0 \\ 0 \end{bmatrix}^T . \tag{6.16}$$

Furthermore, the noise vector W_{k,i_0} and its covariance matrix $\Sigma_{w,d}$ are given by:

$$W_{k,i_0} = \begin{bmatrix} w_{k,i_0} \\ \tilde{w}_k^- \\ \tilde{v}_k^- \end{bmatrix}, \Sigma_{w,d} = \begin{bmatrix} \sigma_w^2 & 0 & 0 \\ 0 & \frac{(N-1)}{N^2}\sigma_w^2 & 0 \\ 0 & 0 & \frac{(N-1)}{N^2}\sigma_v^2 \end{bmatrix}. \tag{6.17}$$

6.4.1 Optimal control using dynamic programming

Let $\hat{x}_{k,i}$ denote the minimum mean square error estimator of $x_{k,i}$ based only on local observations of the i^{th} agent. Also let Y_i^k indicate the column vector of all measurements up to time k at agent i. Next, consider agent i_0 with its cost function given by:

$$J_{i_0} \triangleq \mathbb{E}\left(\sum_{k=0}^{T-1}(x_{k,i_0}^2 + ru_{k,i_0}^2) + x_{T,i_0}^2\right), \tag{6.18}$$

and let V_{k,i_0} be the optimal expected value of J_{i_0} starting from time k on, knowing the measurements $Y_{i_0}^k$ and using an optimal u_{k,i_0}^*, i.e. the optimal cost to go starting from time k. Then using the dynamic programming principle we have:

$$V_{k,i_0} = \min_{u_{k,i_0}} \mathbb{E}\{x_{k,i_0}^2 + ru_{k,i_0}^2 + V_{k+1,i_0}|Y_{i_0}^k\}. \tag{6.19}$$

Now, for tractability, we make the approximation that the optimal cost-to-go is *a quadratic function of the last two most recent measurements*. Recall however that in general, the optimal cost to go will be a quadratic function of all current and past measurements. Therefore, by substituting

$$V_{k+1,i_0} = \begin{bmatrix} y_{k+1,i_0} \\ y_{k,i_0} \end{bmatrix}^T \begin{bmatrix} q_{11,k+1} & q_{12,k+1} \\ q_{12,k+1} & q_{22,k+1} \end{bmatrix} \begin{bmatrix} y_{k+1,i_0} \\ y_{k,i_0} \end{bmatrix} + \bar{q}_{k+1} \tag{6.20}$$

in (6.19) we get:

$$V_{k,i_0} = \min_{u_{k,i_0}} \mathbb{E}\{x_{k,i_0}^2 + ru_{k,i_0}^2 + q_{11,k+1}y_{k+1,i_0}^2 + q_{22,k+1}y_{k,i_0}^2 + 2q_{12,k+1}y_{k,i_0}y_{k+1,i_0} + \bar{q}_{k+1}|Y_{i_0}^k\}, \tag{6.21}$$

which can be written as:

$$\begin{aligned} V_{k,i_0} = &\min_{u_{k,i_0}} \mathbb{E}\{(x_{k,i_0} - \hat{x}_{k,i_0} + \hat{x}_{k,i_0})^2 + ru_{k,i_0}^2 + q_{11,k+1}(y_{k+1,i_0} - \hat{y}_{k+1|k,i_0} + \hat{y}_{k+1|k,i_0})^2 \\ &+ q_{22,k+1}y_{k,i_0}^2 + 2q_{12,k+1}y_{k,i_0}\hat{y}_{k+1|k,i_0} + \bar{q}_{k+1}|Y_{i_0}^k\}, \end{aligned} \tag{6.22}$$

where $\hat{y}_{k+1|k,i_0} = \mathbb{E}\left[y_{k+1,i_0}|Y_{i_0}^k\right]$. Now, using the orthogonality of estimation error with the estimate itself (i.e., $(x_{k,i_0} - \hat{x}_{k,i_0}) \perp \hat{x}_{k,i}$) in (6.22) we get:

$$\begin{aligned} V_{k,i_0} = &\min_{u_{k,i_0}} \mathbb{E}\{(x_{k,i_0} - \hat{x}_{k,i_0})^2 + \hat{x}_{k,i_0}^2 + ru_{k,i_0}^2 + q_{11,k+1}(y_{k+1,i_0} - \hat{y}_{k+1|k,i_0})^2 + q_{11,k+1}\hat{y}_{k+1|k,i_0}^2 \\ &+ q_{22,k+1}y_{k,i_0}^2 + 2q_{12,k+1}y_{k,i_0}\hat{y}_{k+1|k,i_0} + \bar{q}_{k+1}|Y_{i_0}^k\}, \end{aligned} \tag{6.23}$$

which can be expressed as:

$$\begin{aligned} V_{k,i_0} = &\min_{u_{k,i_0}} \mathbb{E}\{(x_{k,i_0} - \hat{x}_{k,i_0})^2 + \hat{x}_{k,i_0}^2 + ru_{k,i_0}^2 + q_{11,k+1}((c + \frac{h}{N})(x_{k+1,i_0} - \hat{x}_{k+1|k,i_0}) \\ &+ h(m_{k+1}^- - \hat{m}_{k+1|k,i_0}^-) + v_{k+1,i_0})^2 + q_{22,k+1}y_{k,i_0}^2 + q_{11,k+1}((c + \frac{h}{N})\hat{x}_{k+1|k,i_0} + h\hat{m}_{k+1|k,i_0}^-)^2 \\ &+ 2q_{12,k+1}y_{k,i_0}((c + \frac{h}{N})\hat{x}_{k+1|k,i_0} + h\hat{m}_{k+1|k,i_0}^-) + \bar{q}_{k+1}|Y_{i_0}^k\}, \end{aligned} \tag{6.24}$$

where $\hat{x}_{k+1|k,i_0} = \mathbb{E}\left[x_{k+1,i_0}|Y_{i_0}^k\right]$ and $\hat{m}_{k+1|k,i_0}^- = \mathbb{E}\left[m_{k+1,i_0}^-|Y_{i_0}^k\right]$. Note here the dependence of the cost to go on the expected state estimation error variance, itself a deterministic function of all current and past state feedback gains in the control law, thus highlighting the dual effects of controls in this context. Then replacing (6.3) and

$$\hat{x}_{k+1|k,i_0} = a\hat{x}_{k,i_0} + bu_{k,i_0} \tag{6.25}$$

in (6.24) and also noting that m_{k+1}^- is considered independent of u_{k,i_0} because of the size of N, we let $\frac{\partial V_{k,i_0}}{\partial u_{k,i_0}} = 0$, which yields:

$$u_{k,i_0}^* = -\frac{b(c + \frac{h}{N})}{r + b^2(c + \frac{h}{N})^2 q_{11,k+1}} \left(a(c + \frac{h}{N})q_{11,k+1}\hat{x}_{k,i_0} + hq_{11,k+1}\hat{m}_{k+1|k,i_0}^- + q_{12,k+1}y_{k,i_0}\right). \tag{6.26}$$

Moreover, using (6.10)-(6.12) the optimal controller u^*_{k,i_0} can be expressed as:

$$u^*_{k,i_0} = -\frac{b(c+\frac{h}{N})}{r+b^2(c+\frac{h}{N})^2 q_{11,k+1}}(q_{11,k+1}\Phi_k^T \hat{X}_{k,i_0} + q_{12,k+1} y_{k,i_0}), \tag{6.27}$$

where

$$\Phi_k = \begin{bmatrix} a(c+\frac{h}{N}) - bf_{1,k}\frac{(N-1)}{N^2}h^2 \\ -bf_{2,k}\frac{(N-1)}{N^2}h^2 \\ [a - bf_{1,k}(c+\frac{(N-1)}{N}h)]h \\ -bf_{2,k}(c+\frac{(N-1)}{N}h)h \\ -bf_{2,k}h \end{bmatrix}. \tag{6.28}$$

Now, we assume that any estimate we make depends at most on the two most recent measurements so that the quadratic assumption about the cost structure remains true from one step to another (see the Kalman filtering section 6.4.2 for more details, as well as the calculation of the filter gains in (6.46)-(6.48)). We then have:

$$\hat{X}_{k,i_0} = \tilde{K}_{1,k} y_{k,i_0} + \tilde{K}_{2,k} y_{k-1,i_0}. \tag{6.29}$$

Thus, u^*_{k,i_0} can further be expressed as follows:

$$u^*_{k,i_0} = -f^*_{1,k} y_{k,i_0} - f^*_{2,k} y_{k-1,i_0}, \tag{6.30}$$

where

$$f^*_{1,k} = \frac{b(c+\frac{h}{N})}{r+b^2(c+\frac{h}{N})^2 q_{11,k+1}}\left(q_{11,k+1}\Phi_k^T \tilde{K}_{1,k} + q_{12,k+1}\right), \tag{6.31}$$

$$f^*_{2,k} = \frac{b(c+\frac{h}{N})}{r+b^2(c+\frac{h}{N})^2 q_{11,k+1}}\left(q_{11,k+1}\Phi_k^T \tilde{K}_{2,k}\right). \tag{6.32}$$

As stated, a necessary condition for the posited feedback sequence to be Nash is that the agent i_0 recovers the feedback gains he assumed optimal in the first place, i.e. the fixed point equilibrium condition holds, which yields:

$$(f_{1,k}, f_{2,k}) = (f^*_{1,k}, f^*_{2,k}). \tag{6.33}$$

Furthermore, plugging in (6.23) the optimal controller u^*_{k,i_0} and using estimate (6.29) we get, after some calculations, the following backward coupled equations:

$$\begin{aligned}q_{11,k} =& q_{11,k+1}(-bf_{1,k}(c+\frac{h}{N})+\Phi_k^T\tilde{K}_{1,k})^2+2q_{12,k+1}(-bf_{1,k}(c+\frac{h}{N})+\Phi_k^T\tilde{K}_{1,k})+q_{22,k+1}\\ &+\tilde{K}_{1,k}^{(x)2}+rf_{1,k}^2,\end{aligned} \tag{6.34}$$

$$q_{22,k} = q_{11,k+1}(-bf_{2,k}(c+\frac{h}{N})+\Phi_k^T\tilde{K}_{2,k})^2+\tilde{K}_{2,k}^{(x)2}+rf_{2,k}^2, \tag{6.35}$$

$$\begin{aligned}q_{12,k} =& \tilde{K}_{1,k}^{(x)}\tilde{K}_{2,k}^{(x)}+q_{12,k+1}(-bf_{2,k}(c+\frac{h}{N})+\Phi_k^T\tilde{K}_{2,k})+rf_{1,k}f_{2,k}\\ &+q_{11,k+1}(-bf_{1,k}(c+\frac{h}{N})+\Phi_k^T\tilde{K}_{1,k})(-bf_{2,k}(c+\frac{h}{N})+\Phi_k^T\tilde{K}_{2,k}),\end{aligned} \tag{6.36}$$

$$\bar{q}_k = q_{11,k+1}(\Phi_k^T P_{k|k}\Phi_k+(c^2+\frac{2ch+h^2}{N})\sigma_w^2+(1+b^2f_{1,k}^2h^2\frac{(N-1)}{N^2})\sigma_v^2)+\bar{q}_{k+1}+P_{k|k}^{(xx)}, \tag{6.37}$$

where $\tilde{K}_{1,k}^{(x)}$, $\tilde{K}_{2,k}^{(x)}$ are respectively, the first elements of the filter gain vectors $\tilde{K}_{1,k}$, $\tilde{K}_{2,k}$. Moreover, $P_{k|k}^{(xx)}$ denotes the first entry in the main diagonal of the estimation error covariance matrix $P_{k|k}$.

6.4.2 Kalman filter-based estimation

Applying the standard (time-varying) Kalman filter algorithm to the dynamics (6.10)-(6.11) we have:

$$\hat{X}_{k+1,i_0} = A_{d,k}\hat{X}_{k,i_0}+B_d u_{k,i_0}+K_{k+1}(y_{k+1,i_0}-H_d(A_{d,k}\hat{X}_{k,i_0}+B_d u_{k,i_0})) \tag{6.38}$$

$$P_{k+1|k} = A_{d,k}P_{k|k}A_{d,k}^T+D_{d,k}\Sigma_{w,d}D_{d,k}^T \tag{6.39}$$

$$K_{k+1} = P_{k+1|k}H_d^T(H_dP_{k+1|k}H_d^T+R_d)^{-1} \tag{6.40}$$

$$P_{k+1|k+1} = (I-K_{k+1}H_d)P_{k+1|k} \tag{6.41}$$

which is initialized via given $\hat{X}_{i,0}$ and $P_{0|-1}$. Also, $R_d=\sigma_v^2$. Then we note that the filtering equation (6.38) can be written as

$$\hat{X}_{k+1,i_0} = (A_{d,k}-K_{k+1}H_dA_{d,k})\hat{X}_{k,i_0}+(I-K_{k+1}H_d)B_du_{k,i_0}+K_{k+1}y_{k+1,i_0} \tag{6.42}$$

Thus, replacing

$$\hat{X}_{k,i_0} = (A_{d,k-1} - K_k H_d A_{d,k-1})\hat{X}_{k-1,i_0} + (I - K_k H_d)B_d u_{k-1,i_0} + K_k y_{k,i_0} \tag{6.43}$$

in (6.42) we get:

$$\begin{aligned}\hat{X}_{k+1,i_0} =& (A_{d,k} - K_{k+1} H_d A_{d,k})[(A_{d,k-1} - K_k H_d A_{d,k-1})\hat{X}_{k-1,i_0} + (I - K_k H_d)B_d u_{k-1,i_0}] \\ &+ (I - K_{k+1} H_d) B_d u_{k,i_0} + (A_{d,k} - K_{k+1} H_d A_{d,k}) K_k y_{k,i_0} + K_{k+1} y_{k+1,i_0} \end{aligned} \tag{6.44}$$

Now, we assume that any estimate we make depends at most on the two most recent measurements so that the quadratic assumption relating to the cost structure remains true from one step to another. Therefore, also letting the optimal $u^*_{k,i_0} = -f_{1,k} y_{k,i_0} - f_{2,k} y_{k-1,i_0}$ in (6.44) we get:

$$\hat{X}_{k+1,i_0} = [-f_{1,k}(I - K_{k+1} H_d) B_d + (A_{d,k} - K_{k+1} H_d A_{d,k}) K_k] y_{k,i_0} + K_{k+1} y_{k+1,i_0}. \tag{6.45}$$

Hence, we have:

$$\hat{X}_{k,i_0} = \tilde{K}_{1,k} y_{k,i_0} + \tilde{K}_{2,k} y_{k-1,i_0}, \tag{6.46}$$

where

$$\tilde{K}_{1,k} = K_k, \tag{6.47}$$

$$\tilde{K}_{2,k} = - f_{1,k-1}(I - K_k H_d) B_d + (A_{d,k-1} - K_k H_d A_{d,k-1}) K_{k-1}. \tag{6.48}$$

6.4.3 Initialization

Over a finite time horizon $[0, T]$, the solution starts with a forward sweep whereby one assumes an initial set of output feedback gains $([f_{1,1}, f_{1,2}, ..., f_{1,T}], [f_{2,1}, f_{2,2}, ..., f_{2,T}])$, which gives an expression of state estimates as well as their error covariances in terms of the assumed gains and the measurements that will be gathered over time, through recursive equations (6.39)-(6.41) and (6.46)-(6.48). Then, by proceeding through a backward sweep, one can use these values to find $q_{11,k}$, $q_{22,k}$, $q_{12,k}$, $\bar{q}_k$, for all k's, through recursive equations (6.34)-(6.37) and hence a new set of candidate output feedback gains. The forward-backward sweep stops whenever one reaches a fixed point in the space of output feedback gains. In particular, after a first Kalman filtering based forward calculations cycle, we initialize the backward sweep calculation at time T as follows:

$$\begin{aligned} V_{T,i_0} &= \mathbb{E}\{x_{T,i_0}^2|Y_{i_0}^T\} \\ &= \mathbb{E}\{(x_{T,i_0} - \hat{x}_{T,i_0} + \hat{x}_{T,i_0})^2|Y_{i_0}^T\} \\ &= \mathbb{E}\{(x_{T,i_0} - \hat{x}_{T,i_0})^2 + \hat{x}_{T,i_0}^2|Y_{i_0}^T\} \\ &= P_{T|T}^{(xx)} + (\tilde{K}_{1,T}^{(x)} y_{T,i_0} + \tilde{K}_{2,T}^{(x)} y_{T-1,i_0})^2 \end{aligned} \tag{6.49}$$

Thus, we get:

$$q_{11,T} = \tilde{K}_{1,T}^{(x)2}, \quad q_{22,T} = \tilde{K}_{2,T}^{(x)2}, \quad q_{12,T} = \tilde{K}_{1,T}^{(x)} \tilde{K}_{2,T}^{(x)}, \quad \bar{q}_T = P_{T|T}^{(xx)}. \tag{6.50}$$

6.5 Numerical results

The numerical results reported in this section are obtained considering the following parameter setting for a representative agent, i.e., the $1st$ agent: $b = c = h = 1$ and $\sigma_w = \sigma_v = 1$, with initial standard deviation $\sigma_0 = 1$ and $\mathbb{E}x_{0,i} = \bar{x}_0 = 0$ for all agents $i = 1, 2, \ldots, N$. In addition, we will only deal with the case $a \geq 0$ (the symmetric property should hold for the $a \leq 0$ case). The value of a and f will be specified in the different simulations and $r = 1$. We experiment first, close to the a stability region $[0, 2.53]$ by Kalman-Riccati couple, and initialize the gains using the steady-state isolated (naive) Kalman sequence (see Appendix for details) and the Riccati control gain f^*, where f^* is obtained based on the positive solution of the associated algebraic Riccati equation

$$b^2\Sigma^2 + (r - a^2 r - b^2)\Sigma - r = 0, \tag{6.51}$$

with

$$f^* = \frac{ab\Sigma}{r + b^2\Sigma} \cdot \tag{6.52}$$

In particular, we initialize the proposed algorithm by combining (6.52), (6.55) and (6.59) with the two most recent measurements, which yields

$$f_1 = f^* K^*, \tag{6.53}$$

$$f_2 = f^*(a - bf^*)(1 - cK^*)K^*. \tag{6.54}$$

Then using a *continuation approach*, we let a go to $a+\Delta a$, and intialize the algorithm with the latest sequence. It is observed that the proposed approach, which is the optimal solution with limited memory, improves the previous value of $a_{sup} = 2.53$, obtained in Abedinpour Fallah *et al.* (2016), up to the new value of 2.72, where a_{sup} is the maximum value of a such that

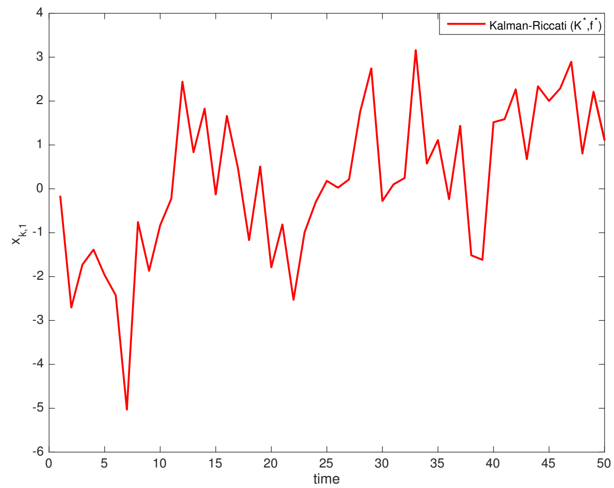

Figure 6.2 Behavior of agent 1 using Kalman-Riccati couple $K^* = 0.809$, $f^* = 1.618$, where $a = 2$, $N = 100$, $\bar{J}^{(N)}/T = 15.65$.

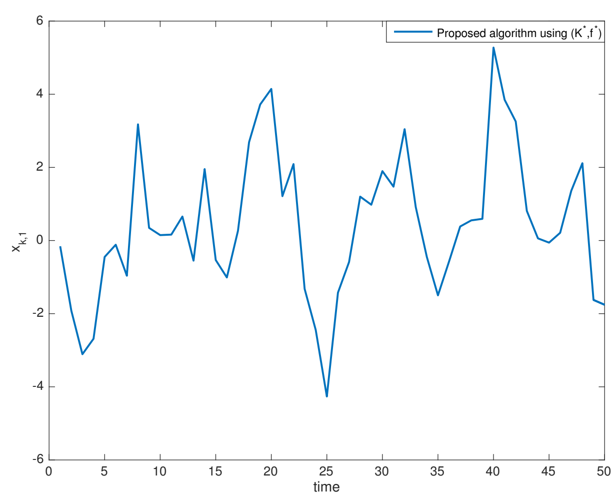

Figure 6.3 Behavior of agent 1 using the proposed algorithm which is initialized by Kalman-Riccati couple $K^* = 0.809$, $f^* = 1.618$, where $a = 2$, $N = 100$, $\bar{J}^{(N)}/T = 15.08$.

the naive optimal Kalman-Riccati couple (K^*, f^*) is inside of the stability region and can stabilize the population (see section 4.4.2. in Abedinpour Fallah *et al.* (2016)). Moreover, the behaviors of agent 1 using the Kalman-Riccati couple versus the proposed algorithm are compared in Figs. 6.2-6.3 and Figs. 6.4-6.5 respectively, for $a = 2$ and $a = 2.6$, where $\bar{J}^{(N)}$ denotes the LQ average cost over all the population, that is, $\bar{J}^{(N)} = \frac{1}{N}\sum_{j=1}^{N} J_j$. Furthermore, Figs. 6.6-6.7 depict the behaviors of agent 1 using the Kalman gain with $a_f = 0.9$ (where $a_f = a - bf$) versus the proposed algorithm for $a = 2.6$.

In addition, we illustrate the sequence of output feedback gains $(f_{1,k}, f_{2,k})$ that were obtained in every case. In particular, for Figure 6.2 with $a = 2$ we have:

$[f_{1,1}, f_{1,2}, ..., f_{1,50}] = [1.309, 1.309, ..., 1.309]$,
$[f_{2,1}, f_{2,2}, ..., f_{2,50}] = [0.0955, 0.0955, ..., 0.0955]$,

which are obtained from (6.53)-(6.54). While for Figure 6.3 with $a = 2$, from the proposed algorithm we have:

$[f_{1,1}, f_{1,2}, ..., f_{1,50}] = [1.1797, 1.2532, 1.2445, 1.2444, 1.2442, 1.2435, 1.2404, 1.2276, 1.1688, 0.6857]$,

$[f_{2,1}, f_{2,2}, ..., f_{2,50}] = [0, 0.1338, 0.1196, 0.1195, 0.1193, 0.1180, 0.1126, 0.0663]$.

Similarly, for Figure 6.4 with $a = 2.6$ we have:

$[f_{1,1}, f_{1,2}, ..., f_{1,50}] = [1.9838, 1.9838, ..., 1.9838]$,

$[f_{2,1}, f_{2,2}, ..., f_{2,50}] = [0.0825, 0.0825, ..., 0.0825]$,

while for Figure 6.5 with $a = 2.6$ we have:

$[f_{1,1}, f_{1,2}, ..., f_{1,50}] = [1.7997, 1.9086, 1.8895, 1.8916, 1.8913, 1.8914, 1.8913, 1.8913, 1.8913, 1.8913, 1.8913, 1.8913, 1.8911, 1.8914, 1.8909, 1.8912, 1.8891, 1.8948, 1.8705, 1.8411, 1.0163]$,

$[f_{2,1}, f_{2,2}, ..., f_{2,50}] = [0, 0.2280, 0.1964, 0.1988, 0.1980, 0.1981, 0.1982, 0.1982, 0.1986, 0.1980, 0.1992, 0.1986, 0.2016, 0.1909, 0.2220, 0.1831, 0.0640]$.

Moreover, for Figure 6.6 and Figure 6.7 with the same $a = 2.6$, respectively, we have:

$[f_{1,1}, f_{1,2}, ..., f_{1,50}] = [1.4849, 1.4849, ..., 1.4849]$,

$[f_{2,1}, f_{2,2}, ..., f_{2,50}] = [0.1691, 0.1691, ..., 0.1691]$,

and

$[f_{1,1}, f_{1,2}, ..., f_{1,50}] = [1.7918, 1.9039, 1.8942, 1.8971, 1.8960, 1.8963, 1.8964, 1.8963, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8964, 1.8963, 1.8963, 1.8963, 1.8963, 1.8962, 1.8962, 1.8961, 1.8960, 1.8960, 1.8958, 1.8957, 1.8955, 1.8953, 1.8951, 1.8948, 1.8945, 1.8942, 1.8938, 1.8933, 1.8928, 1.8923, 1.8917, 1.8910, 1.8902, 1.8892, 1.8884, 1.8875, 1.8820, 1.8825, 1.8563, 1.0179]$,

$[f_{2,1}, f_{2,2}, ..., f_{2,50}] = [0, 0.2272, 0.1787, 0.1912, 0.1914, 0.1904, 0.1908, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1907, 0.1908, 0.1908, 0.1908, 0.1909, 0.1909, 0.1910, 0.1911, 0.1912, 0.1914, 0.1915, 0.1917, 0.1920, 0.1923, 0.1926, 0.1930, 0.1935, 0.1940, 0.1946, 0.1953, 0.1960, 0.1968, 0.1976, 0.1985, 0.1997, 0.2011, 0.2025, 0.2037, 0.2053, 0.2106, 0.1962, 0.1651, 0.0671]$.

It is noted that up to the fixed point numerical convergence criterion, the fixed points associated to Figure 6.5 and Figure 6.7 appear to be essentially equivalent, although the two calculations were initialized differently.

6.6 Conclusion

In this paper, motivated by the application of CDMA based communication and control system modeled as an interference induced game in a multi-agent networked control system, we considered the computation of a symmetric Nash equilibrium within a restricted class of output feedback policies. In particular, we explored necessary conditions (fixed point equations) that would lead to Nash equilibria under a class of total measurements preserving time varying output feedback stabilizing controllers, as N goes to infinity, for a finite length 2 time horizon. Subsequently the corresponding expressions were written for horizon larger than 2, by preserving only the contribution of the most recent 2 measurements. This led to a heuristic estimate of Nash equilibria under a class of output time varying controllers with only

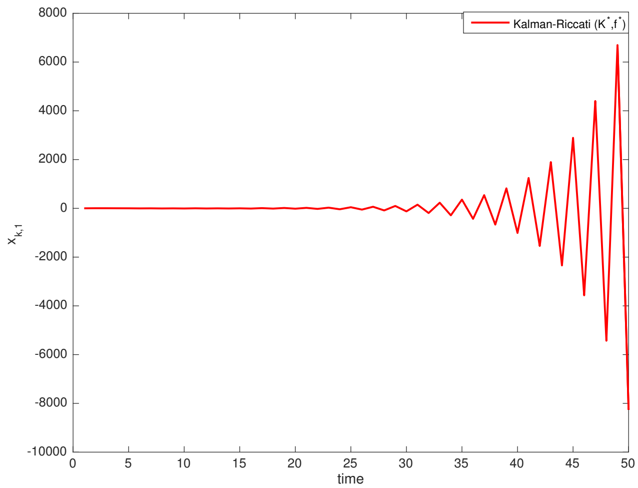

Figure 6.4 Unstable behavior of agent 1 using Kalman-Riccati couple $K^* = 0.8735$, $f^* = 2.2711$, where $a = 2.6$, $N = 100$, $\bar{J}^{(N)}/T = 6.4582 \times 10^7$.

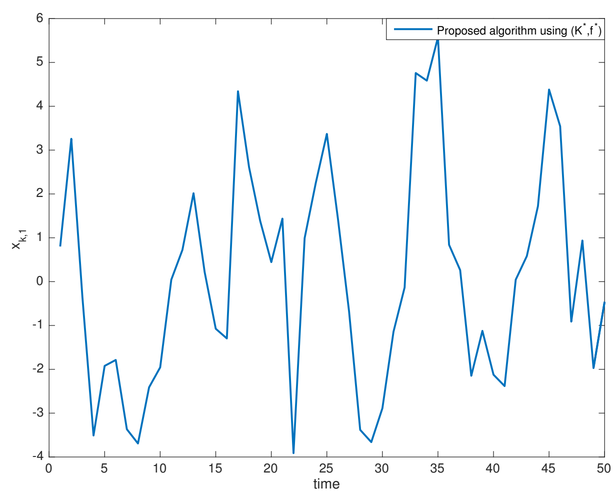

Figure 6.5 Behavior of agent 1 using the proposed algorithm which is initialized by Kalman-Riccati couple $K^* = 0.8735$, $f^* = 2.2711$, where $a = 2.6$, $N = 100$, $\bar{J}^{(N)}/T = 45.6167$.

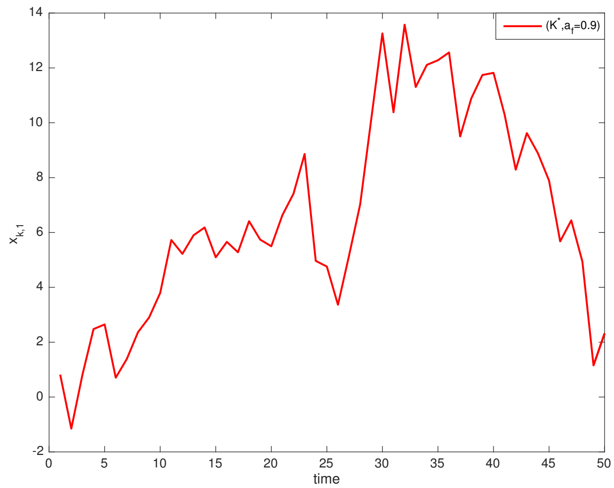

Figure 6.6 Behavior of agent 1 using the Kalman gain $K^* = 0.8735$ and $a_f = 0.9$, where $a = 2.6$, $N = 100$, $\bar{J}^{(N)}/T = 151$.

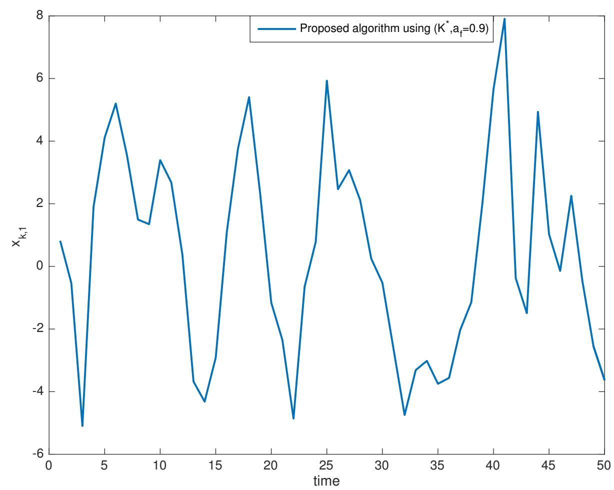

Figure 6.7 Behavior of agent 1 using the proposed algorithm which is initialized by the Kalman gain $K^* = 0.8735$, and $a_f = 0.9$, where $a = 2.6$, $N = 100$, $\bar{J}^{(N)}/T = 44.75$.

recollection of the two most recent measurements. We have shown that for some instances, the range of Nash equilibria obtained from the Kalman-Riccati couple can be extended using the proposed methodology which involves limited measurements memory. In future work, we shall investigate the existence properties of the Nash equilibria that can be achieved within this set up.

6.7 Appendix: The steady-state isolated Kalman sequence

Definition 6.1. *The isolated (naive) Kalman filter is a Luenberger like observer equation under the assumed state estimate feedback structure*

$$u_{k,i} = -f\hat{x}_{k,i}, \tag{6.55}$$

and assuming zero interference in the local measurements (setting $h = 0$ in (6.4)), i.e., it evolves according to:

$$\hat{x}_{k+1,i} = (a - bf)\hat{x}_{k,i} + K^*(y_{k+1,i} - c(a - bf)\hat{x}_{k,i}), \tag{6.56}$$

with the steady-state scalar gain

$$K^* = \frac{cP_\infty}{c^2P_\infty + \sigma_v^2}\,, \tag{6.57}$$

where P_∞ is the unique positive solution of

$$c^2P_\infty^2 + ((1 - a^2)\sigma_v^2 - c^2\sigma_w^2)P_\infty - \sigma_w^2\sigma_v^2 = 0. \tag{6.58}$$

Proposition 6.1. *In the stability region, the isolated Kalman filter equivalent equation in the steady-state is given by*

$$\hat{x}_{k,i} = K_1'y_{k,i} + K_2'y_{k-1,i} + ... + K_k'y_{1,i}, \tag{6.59}$$

where

$$K_j' = (a - bf)^{j-1}(1 - cK^*)^{j-1}K^*. \tag{6.60}$$

Proof: By rearranging (6.56) we have:

$$\hat{x}_{k+1,i} = (a - bf)(1 - cK^*)\hat{x}_{k,i} + K^*y_{k+1,i}. \tag{6.61}$$

Then substituting

$$\hat{x}_{k,i} = K_1'y_{k,i} + K_2'y_{k-1,i} + ... + K_k'y_{1,i}, \tag{6.62}$$

and

$$\hat{x}_{k+1,i} = K_1' y_{k+1,i} + K_2' y_{k,i} + ... + K_{k+1}' y_{1,i}, \tag{6.63}$$

into (6.61) and applying the stationarity property by making the left-hand-side of the resulting equation equal to its right-hand-side, we get the fixed-point values (6.60). $\square$

CHAPTER 7 GENERAL DISCUSSION

In this chapter, we consider discussing the following salient aspects.

7.1 Stability in Mean field games

Mean field games constitute a class of non cooperative stochastic dynamic games, where there is a large number of players or agents, who interact with each other through a mean field coupling term (also known as the mass behavior, which describes the average of all the agents' states) included in the individual cost functions and/or each agent's dynamics (see Huang *et al.* (2007), for example). As a solution concept in non cooperative games theory, one objective of the study of the mean field games is to obtain a characterization of Nash equilibria under the assumption of rationality for each player. In general, however, such a characterization is challenging and difficult because the complexity increases with the number of players and the dimension of the state space. Moreover, it is more realistic that every agent is able to access only his own state information, in which case the conventional centralized dynamic programming approach as in the standard derivation of dynamic Nash equilibria in Başar and Olsder (1998) cannot be applied. In Huang *et al.* (2007), an approximation scheme was developed to estimate the actual mass behavior, which therefore provides an approximated equilibrium solution to the mean field game. In particular, under the Nash certainty equivalence (NCE) principle, the mean field game was analyzed through two steps: i) solving a generic optimal control problem by replacing the mean field coupling term with an arbitrary deterministic function, and ii) approximation of the mean field term by a fixed-point analysis.

In this thesis, we have studied a somewhat dual situation whereby large populations of partially observed stochastic agents, although a priori individually independent, are coupled only via their observation structure. The latter involves an interference term depending on the empirical mean of all agent states. Moreover, in Abedinpour Fallah *et al.* (2013b), for the continuous-time version of the model in this thesis, we have used the state aggregation technique to anticipate the mean field coupling term (see Appendix A). We have learned that assuming independence of agents and also assuming separation between state estimation and control, even for large N, leads to poor local state estimation properties, and as a consequence to poor control performance or even stabilization abilities for initially unstable agent systems. This is because of persistent correlations between agent states by virtue of the coupling in their state estimates which directly feed into their dynamics. The apparent role of individual agent systems lack of stability in interference coupled systems, points at a potential hitherto unsuspected role of instability in more classical mean field games where the mean agent state enters individual dynamics (see Huang *et al.* (2005), for example).

7.2 Finding the best linear control strategies for decentralized LQG systems

As noted by Mahajan and Nayyar (2015), the problem of finding the best linear control strategies for decentralized LQG systems has the following prominent characteristics:

(i) In general, linear control strategies are not globally optimal, i.e., there may exist non linear control strategies that outperform linear strategies. To cite an instance, Witsenhausen (1968) gave a famous counterexample of a simple two-player decentralized LQG problem whose optimal solution is nonlinear.

(ii) In general, the problem of finding the best linear control strategies is not convex.

(iii) In general, the best linear control strategy may not have a finite-dimensional sufficient statistic. More specifically, it may not be possible to represent the best linear controller by a finite set of estimates that are generated by recursions of finite order. For example, see the two-controller completely decentralized system considered in Whittle and Rudge (1974).

In this thesis, we showed that, in general, the Riccati gain f^* is not asymptotically optimal as the number N of agents goes to infinity when using the Bulk Filter. In particular, in the range of a's for which (k^*, f) is stabilizing, the coupling term of the mean asymptotically disappears from the measurements. In that case, f^* is the optimum provided that f^* is in the range of (k^*, f) couples which still stabilize the system, and indeed the bulk filter becomes in that case equivalent to the naive Kalman filter. However, outside of the stationarization range, the quality of estimation depends on the applied feedback gain (dual effect of the control, see Feldbaum (1960); Witsenhausen (1968)) and this explains the non-optimality of the Riccati gain in general. Moreover, it turns out that the bulk filter developed in this thesis, is growing memory while no finite dimensional sufficient statistic appears within grasp.

CHAPTER 8 CONCLUSION AND RECOMMENDATIONS

In this thesis, motivated by the application of code division multiple access (CDMA) based communication and control system modeled as an interference induced game in a multi-agent networked control system framework, we have studied a system of uniform agents coupled via their distinct sets of partial observations, whereby each agent has noisy measurements of its own state. The dissertation is presented in three main parts. Our first interest has been to establish that for certain classes of cost and dynamic parameters, optimal separated control laws obtained by ignoring the interference coupling, are asymptotically optimal when the number of agents is sufficiently high. Moreover, there is a lack of stability threshold past which, the only choice left for the majority of agents is to act cooperatively. The second part has been focused on the extension of the estimation framework in the first part to utilize the exact decentralized filtering under a class of time invariant certainty equivalent feedback controllers. We have numerically observed that the proposed estimator in combination with an arbitrary (stabilizing under perfect state observations) state estimate feedback gain, succeeds in maintaining the boundedness of the closed loop system even when individual systems are highly unstable. Furthermore, since the exact optimal filter is growing dimension, we have devised approximate finite dimensional filters to reduce memory requirements. The final part presented a fixed point based algorithm for identifying Nash equilibria within a restricted class of output feedback policies, where unlike the standard LQG solution, *the estimation error cost depends on the chosen feedback gains*, and thus the structure of the fixed point equations one needs to satisfy also does.

We now conclude by outlining some recommendations and possible future research directions.

- Establishing stability properties of optimal filtered separated feedback laws.
- Understanding the dynamic properties, in particular periodicities, or chaotic behavior of optimal bulk filters.
- Establishing if linear time varying output feedback is a sufficient class for best responses in our system.
- Exploring the multidimensional situation.

— Further investigating the impact of individual agents systems instability on the stability in classical mean field LQG games, when the overall mean state term enters into individual agent dynamics.

— Investigating necessary and sufficient conditions for stability of networked multi-agent control systems subject to time-varying delays and data packet losses.

— Investigating the application of the proposed methodologies to power control of optical networks using a game-theoretic model introduced by Pavel (2006), where a link level power control scheme adjusts the optical signal-to-noise ratio (OSNR) value of the signals toward channel OSNR optimization.

REFERENCES

ABEDINPOUR FALLAH, M., MALHAMÉ, R. and MARTINELLI, F. (2013a). Decentralized estimation in a class of measurements induced mean field control problems. *Proc. of the 52nd IEEE conference on decision and control*. 3146–3151.

ABEDINPOUR FALLAH, M., MALHAMÉ, R. and MARTINELLI, F. (2013b). Distributed estimation and control for large population stochastic multi-agent systems with coupling in the measurements. *Proc. of the european control conference*. 4353–4358.

ABEDINPOUR FALLAH, M., MALHAMÉ, R. and MARTINELLI, F. (2014). Distributed estimation of multi-agent systems with coupling in the measurements: Bulk algorithm and approximate kalman-type filtering. *Proc. of the 53rd IEEE conference on decision and control*. 1810–1815.

ABEDINPOUR FALLAH, M., MALHAMÉ, R. and MARTINELLI, F. (2016). A class of interference induced games: Asymptotic Nash equilibria and parameterized cooperative solutions. *Automatica*, 69, 181–194.

ADLAKHA, S., JOHARI, R., WEINTRAUB, G. and GOLDSMITH, A. (2008). Oblivious equilibrium for large-scale stochastic games with unbounded costs. *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, 5531–5538.

ALTMAN, E., BOULOGNE, T., EL-AZOUZI, R., JIMÉNEZ, T. and WYNTER, L. (2006). A survey on networking games in telecommunications. *Computers and Operations Research*, 33, 286–311.

ANDERSON, B. D. O. and MOORE, J. B. (1979). Optimal filtering. *Englewood Cliffs*, 21, 22–95.

AUMANN, R. (1964). Markets with a continuum of traders. *Econometrica: Journal of the Econometric Society*, 39–50.

AZIZ, M. and CAINES, P. (2014). Computational investigations of decentralized cellular network optimization via mean field control. *Proc. of the 53rd IEEE conference on decision and control*. 5560–5567.

AZIZ, M. and CAINES, P. E. (2017). A mean field game computational methodology for decentralized cellular network optimization. *IEEE Transactions on Control Systems Technology*, 25 (2), 563–576.

BAGCHI, A. (1993). *Optimal control of stochastic systems.* Prentice-Hall, Inc.

BAŞAR, T. and OLSDER, G. (1998). *Dynamic noncooperative game theory.* SIAM.

BUCKDAHN, R., CARDALIAGUET, P. and QUINCAMPOIX, M. (2011). Some recent aspects of differential game theory. *Dynamic Games and Applications*, 1, 74–114.

CAINES, P. (2015). Mean field games. *Encyclopedia of Systems and Control, T. Samad and J. Baillieul, Eds.*, Springer London. 706–712.

CAINES, P. and KIZILKALE, A. (2013). Recursive estimation of common partially observed disturbances in mfg systems with application to large scale power markets. *Proc. of the 52nd IEEE conference on decision and control.* 2505–2512.

CAINES, P. and KIZILKALE, A. (2014). Mean field estimation for partially observed lqg systems with major and minor agents. *In Proc. of the 19th IFAC world congress.*

CAINES, P. E. and KIZILKALE, A. C. (2016). ϵ-Nash equilibria for partially observed lqg mean field games with major player. *IEEE Transactions on Automatic Control*, 1–10.

CARMONA, G. (2004). Nash equilibria of games with a continuum of players.

CHONG, C. and KUMAR, S. (2003). Sensor networks: evolution, opportunities, and challenges. *Proceedings of the IEEE*, 91, 1247–1256.

EKSIN, C., MOLAVI, P., RIBEIRO, A. and JADBABAIE, A. (2014). Bayesian quadratic network game filters. *IEEE Transactions on Signal Processing*, 62, 2250–2264.

FELDBAUM, A. A. (1960). Dual control theory i-iv. *Autom. Remote Control*, 21, 874–880.

FIROOZI, D. and CAINES, P. E. (2016). Mean field game ε-Nash equilibria for partially observed optimal execution problems in finance. *Decision and Control (CDC), 2016 IEEE 55th Conference on.* IEEE, 268–275.

GARIN, F. and SCHENATO, L. (2010). A survey on distributed estimation and control applications using linear consensus algorithms. *Networked Control Systems*, Springer. 75–107.

GELB, A. (1974). *Applied optimal estimation.* MIT press.

GOMES, D., LAFLECHE, L. and NURBEKYAN, L. (2016). A mean-field game economic growth model. *American Control Conference (ACC), 2016.* IEEE, 4693–4698.

GOMES, D. A. and SAÚDE, J. (2014). Mean field games models—a brief survey. *Dynamic Games and Applications*, 4, 110–154.

GRAMMATICO, S. (2016). Exponentially convergent decentralized charging control for large populations of plug-in electric vehicles. *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 5775–5780.

GRÖNBAUM, D. and OKUBO, A. (1994). Modelling social animal aggregations. *Frontiers in theoretical biology*, New York: Springer-Verlag. 296–325.

HELBING, D., FARKAS, I. and VICSEK, T. (2000). Simulating dynamical features of escape panic. *Nature*, 407, 487–490.

HUANG, M., CAINES, P. and MALHAMÉ, R. (2003). Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*. IEEE, vol. 1, 98–103.

HUANG, M., CAINES, P. and MALHAMÉ, R. (2004). Uplink power adjustment in wireless communication systems: a stochastic control analysis. *IEEE Transactions on Automatic Control*, 49, 1693–1708.

HUANG, M., CAINES, P. and MALHAMÉ, R. (2006a). Distributed multi-agent decision-making with partial observations: asymptotic Nash equilibria. *Proc. of the 17th international symposium on mathematical theory of networks and systems, Kyoto, Japan*. 2725–2730.

HUANG, M., CAINES, P. and MALHAMÉ, R. (2006b). Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6, 221–252.

HUANG, M., CAINES, P. and MALHAMÉ, R. (2007). Large-population cost-coupled lqg problems with non-uniform agents: individual-mass behavior and decentralized epsilon-Nash equilibria. *IEEE Transactions on Automatic Control*, 52, 1560–1571.

HUANG, M., CAINES, P. and MALHAMÉ, R. (2012). Social optima in mean field lqg control: centralized and decentralized strategies. *IEEE Transactions on Automatic Control*, 57, 1736–1751.

HUANG, M., MALHAMÉ, R. and CAINES, P. (2005). Nash equilibria for large-population linear stochastic systems of weakly coupled agents. *Analysis, Control and Optimization of Complex Dynamic Systems, edited by E.K. Boukas and R.P. Malhamé*, Springer. 215–252.

KARFOPOULOS, E. and HATZIARGYRIOU, N. (2013). A multi-agent system for controlled charging of a large population of electric vehicles. *IEEE Transactions on Power Systems*, 28, 1196–1204.

KHAN, M. and SUN, Y. (2002). Non-cooperative games with many players. *Handbook of game theory with economic applications*, 3, 1761–1808.

KIAN, A. R., JR., J. B. C. and SIMAAN, M. A. (2002). Stochastic discrete-time Nash games with constrained state estimators. *Journal of optimization theory and applications*, 114, 171–188.

KIM, J. and ROCK, S. (2006). Stochastic feedback controller design considering the dual effect. *Proceedings of the AIAA Guidance, Navigation and Control Conference*.

KIZILKALE, A. C. and MALHAMÉ, R. P. (2014). Collective target tracking mean field control for markovian jump-driven models of electric water heating loads. *IFAC Proceedings Volumes*, 47, 1867–1872.

KIZILKALE, A. C. and MALHAMÉ, R. P. (2016). Collective target tracking mean field control for markovian jump-driven models of electric water heating loads. *Control of Complex Systems: Theory and Applications, Edited by K. G. Vamvoudakis and S. Jagannathan*, Butterworth-Heinemann: Elsevier. 559–584.

KOSKIE, S. and GAJIC, Z. (2006). Signal-to-interference-based power control for wireless networks: A survey, 1992–2005. *Dynamics of Continuous, Discrete and Impulsive Systems B: Applications and Algorithms*, 13, 187–220.

KUMAR, P. R. and VARAIYA, P. (1986). *Stochastic systems: estimation, identification and adaptive control*. Prentice-Hall, Inc.

LACHAPELLE, A. (2010). Human crowds and groups interactions: a mean field games approach. *Technical report, CEREMADE, Université Paris-Dauphine*.

LASRY, J. and LIONS, P. (2006a). Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343, 619–625.

LASRY, J. and LIONS, P. (2006b). Jeux à champ moyen. ii–horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343, 679–684.

LASRY, J. and LIONS, P. (2007). Mean field games. *Japanese Journal of Mathematics*, 2, 229–260.

LE NY, J., FERON, E. and DAHLEH, M. (2011). Scheduling continuous-time kalman filters. *IEEE Transactions on Automatic Control*, 56, 1381–1394.

LI, T. and ZHANG, J. F. (2008). Asymptotically optimal decentralized control for large population stochastic multiagent systems. *IEEE Transactions on Automatic Control*, 53, 1643–1660.

LIAN, F., CHAKRABORTTY, A. and DUEL-HALLEN, A. (2017). Game-theoretic multi-agent control and network cost allocation under communication constraints. *IEEE Journal on Selected Areas in Communications*, PP, 1–1.

MA, Z., CALLAWAY, D. and HISKENS, I. (2013). Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Transactions on Control Systems Technology*, 21, 67–78.

MAHAJAN, A. and NAYYAR, A. (2015). Sufficient statistics for linear control strategies in decentralized systems with partial history sharing. *IEEE Transactions on Automatic Control*, 60, 2046–2056.

MANOLAKIS, D. G., INGLE, V. K. and KOGON, S. M. (2005). *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*, vol. 46. Artech House Norwood.

MAO, X. and YUAN, C. (2006). *Stochastic differential equations with Markovian switching*. Imperial College Press.

MOON, J. and BAŞAR, T. (2014). Discrete-time lqg mean field games with unreliable communication. *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2697–2702.

NEUMANN, J. V. and MORGENSTERN, O. (1953). *Theory of games and economic behavior*. Princeton university press, 3rd edition.

NOBLE, B. and DANIEL, J. W. (1988). *Applied linear algebra*. Prentice Hall, Englewood Cliffs, NJ.

NOURIAN, M., CAINES, P., MALHAMÉ, R. and HUANG, M. (2012). Mean field lqg control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation. *IEEE Transactions on Automatic Control*, 57, 2801–2816.

OGATA, K. (1995). *Discrete-time control systems*. Prentice-Hall Englewood Cliffs, NJ.

OLFATI-SABER, R. and JALALKAMALI, P. (2012). Coupled distributed estimation and control for mobile sensor networks. *IEEE Transactions on Automatic Control*, 57, 2609–2614.

PARISE, F., COLOMBINO, M., GRAMMATICO, S. and LYGEROS, J. (2014). Mean field constrained charging policy for large populations of plug-in electric vehicles. *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 5101–5106.

PASSINO, K. (2002). Biomimicry of bacterial foraging for distributed optimization and control. *IEEE Control Systems Magazine*, 22, 52–67.

PAVEL, L. (2006). A noncooperative game approach to osnr optimization in optical networks. *IEEE Transactions on Automatic Control*, 51, 848–852.

PEQUITO, S., AGUIAR, A., SINOPOLI, B. and GOMES, D. (2011). Nonlinear estimation using mean field games. *Network Games, Control and Optimization (NetGCooP), 2011 5th International Conference on*. IEEE, 1–5.

PERREAU, S. and ANDERSON, M. (2006). A new method for centralised and decentralised robust power control in cdma systems. *Digital Signal Processing*, 16, 568–576.

REIF, J. H. and WANG, H. Y. (1999). Social potential fields: A distributed behavioral control for autonomous robots. *Robotics and Autonomous Systems*, 27, 171–194.

REIF, K., GÜNTHER, S., YAZ, E. and UNBEHAUEN, R. (2000). Stochastic stability of the continuous-time extended kalman filter. *IEE Proceedings-Control Theory and Applications*, 147, 45–52.

RHODES, I. B. and LUENBERGER, D. G. (1969a). Differential games with imperfect state information. *IEEE Transactions on Automatic Control*, 14, 29–38.

RHODES, I. B. and LUENBERGER, D. G. (1969b). Stochastic differential games with constrained state estimators. *IEEE Transactions on Automatic Control*, 14, 476–481.

RITZERFELD, J. H. F. (2005). On stability tests for continuous and discrete-time linear systems. *Proc. of the 16th Annual Workshop on Circuits, Systems and Signal Processing*. 668–673.

ROSHANY-YAMCHI, S., CYCHOWSKI, M., NEGENBORN, R., SCHUTTER, B. D., DELANEY, K. and CONNELL, J. (2013). Kalman filter-based distributed predictive control of large-scale multi-rate systems: Application to power networks. *IEEE Transactions on Control Systems Technology*, 21, 27–39.

RUESS, J., MILIAS-ARGEITIS, A., SUMMERS, S. and LYGEROS, J. (2011). Moment estimation for chemically reacting systems by extended kalman filtering. *The Journal of chemical physics*, 135, 10B621.

SAKSENA, V. and CRUZ, J. (1982). A multimodel approach to stochastic Nash games. *Automatica*, 18, 295–305.

SALEHISADAGHIANI, F. and PAVEL, L. (2016). Distributed Nash equilibrium seeking: A gossip-based algorithm. *Automatica*, 72, 209–216.

ŞEN, N. and CAINES, P. E. (2016). Mean field game theory for agents with individual-state partial observations. *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 6105–6110.

TEMBINE, H., LASAULCE, S. and JUNGERS, M. (2010). Joint power control-allocation for green cognitive wireless networks using mean field theory. *Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM), 2010 Proceedings of the Fifth International Conference on*. IEEE, 1–5.

TSE, D. and HANLY, S. (1999). Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Transactions on Information Theory*, 45, 641–657.

VERDÚ, S. and SHAMAI, S. (1999). Spectral efficiency of cdma with random spreading. *IEEE Transactions on Information Theory*, 45, 622–640.

WANG, B. and ZHANG, J. (2013). Distributed output feedback control of markov jump multi-agent systems. *Automatica*, 49, 1397–1402.

WANG, B. C. and ZHANG, J. F. (2012). Mean field games for large-population multiagent systems with markov jump parameters. *SIAM Journal on Control and Optimization*, 50, 2308–2334.

WANG, B.-C. and ZHANG, J.-F. (2014). Hierarchical mean field games for multi-agent systems with tracking-type costs: distributed epsilon-stackelberg equilibria. *IEEE Transactions on Automatic Control*, 59, 2241–2247.

WEINTRAUB, G. Y., BENKARD, C. L. and ROY, B. V. (2008). Markov perfect industry dynamics with many firms. *Econometrica*, 76, 1375–1411.

WHITTLE, P. and RUDGE, J. (1974). The optimal linear solution of a symmetric team control problem. *Journal of Applied Probability*, 11, 377–381.

WITSENHAUSEN, H. S. (1968). A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6, 131–147.

XIE, L. and KHARGONEKAR, P. (2012). Lyapunov-based adaptive state estimation for a class of nonlinear stochastic systems. *Automatica*, 48, 1423–1431.

YANG, P., FREEMAN, R. and LYNCH, K. (2008). Multi-agent coordination by decentralized estimation and control. *IEEE Transactions on Automatic Control*, 53, 2480–2496.

YANG, T., MEHTA, P. G. and MEYN, S. P. (2011a). Feedback particle filter with mean-field coupling. *Proc. of the 50th IEEE conference on decision and control and european control conference*. 7909–7916.

YANG, T., MEHTA, P. G. and MEYN, S. P. (2011b). A mean-field control-oriented approach to particle filtering. *American Control Conference (ACC), 2011*. IEEE, 2037–2043.

YIN, H., MEHTA, P. G., MEYN, S. P. and SHANBHAG, U. V. (2012). Synchronization of coupled oscillators is a game. *IEEE Transactions on Automatic Control*, 57, 920–935.

YÜKSEL, S. and BAŞAR, T. (2013). *Stochastic networked control systems: Stabilization and optimization under information constraints*. Springer Science & Business Media.

ZHU, Z., LAMBOTHARAN, S., CHIN, W. and FAN, Z. (2016). A mean field game theoretic approach to electric vehicles charging. *IEEE Access*, 4, 3501–3510.

APPENDIX A

ARTICLE 4 : DISTRIBUTED ESTIMATION AND CONTROL FOR LARGE POPULATION STOCHASTIC MULTI-AGENT SYSTEMS WITH COUPLING IN THE MEASUREMENTS

Mehdi Abedinpour Fallah, Roland P. Malhamé, and Francesco Martinelli
European control conference (ECC), pp. 4353-4358, 2013.

A.1 Abstract

In this paper, we investigate a class of large population stochastic multi-agent systems where the agents have linear stochastic dynamics and are coupled via their measurement equations. Using the state aggregation technique, we propose a distributed estimation and control algorithm that combines the Kalman filtering for state estimation and the linear-quadratic-Gaussian (LQG) feedback controller. Moreover, the stability analysis in terms of exponential boundedness in the mean square is given for the proposed algorithm.

A.2 Introduction

In recent years, analysis and control design for large population stochastic multi-agent systems have become an active area in the study and control of complex systems (Huang *et al.* (2006a); Huang *et al.* (2012)). Many practical applications and examples of these systems arise in engineering, biological, social and economic fields (Chong and Kumar (2003); Lachapelle (2010)).

In conventional control systems, control laws are constructed based upon the overall states of the plants. However, in complex systems with many agents, each agent has a self-governed but limited capability of sensing, decision-making and communication. Therefore an important issue is the development of decentralized solutions so that each individual agent may implement a strategy based on its local information together with statistical information on the population of agents. Just as stabilization and optimization are two fundamental issues for single-agent systems, for large population stochastic multi-agent systems we are also concerned with how to construct decentralized control laws that preserve closed-loop system stability while optimizing the performance of agents in a cooperative or non-cooperative (the focus of this paper) context.

Agent to agent interaction during competitive decision-making is usually due to the coupling in their dynamics or cost functions. Specifically, the dynamic coupling is used to specify an environment effect to the individual's decision-making generated by the population of other agents. While each agent only receives a negligible influence from any other given individual, the overall effect of the population (i.e., the mass effect) is significant for each agent's strategy selection.

The state estimation problem has been a fundamental and a challenging problem in theory and applications of control systems. A new formulation of particle filters inspired by the mean-field game theoretic framework of Huang *et al.* (2007), was presented in Yang *et al.* (2011b), Yang *et al.* (2011a). Mean field based distributed multi-agent decision-making with partial observation was studied in Huang *et al.* (2006a), where the considered agents were weakly coupled through both individual dynamics and costs. In this paper, we study a somewhat dual situation whereby large populations of partially observed stochastic agents, although a priori individually independent, are coupled only via their observation structure. More specifically, the "quality" of individual state measurements is affected by certain statistics of the rest of agent states, such as mean, variance, and in the most general case, the instantaneous empirical distribution of these states. It is the latter which in the limit of an infinite population is referred to as the mean field.

Individual agent dynamics are assumed to be linear, stochastic, with linear local state measurements, and in the current paper, the measurements interaction model is assumed to depend only on the empirical mean of agents states, either in a purely additive manner or through the variance of the local measurement. Each agent is associated with an exponentially discounted individual quadratic cost function, and we look for possible, mean field based, Nash equilibrium inducing decentralized control laws as the number of agents grows without bounds.

The study of such measurement-coupled systems is inspired by a variety of applications, for instance the communications model for power control in cellular telephone systems (Huang *et al.* (2004)), where the received signal of a given user at the base station views all other incell user signals, as well as other cell signals arriving at the base station, as interference or noise. In general, the model is aimed at capturing various forms of interference effects from the environment on individual agent observations.

A.3 System Model and Problem Statement

Consider a system of n agents, each with scalar dynamics where the evolution of the state component is described by

$$dz_i = (az_i + bu_i)dt + \sigma_w dw_i \tag{A.1}$$

and the evolution of the measured output is given by either of the following

— Case (a):

$$dy_i = cz_i dt + (\sigma_v + h\left(\frac{1}{n}\sum_{j=1}^{n} z_j\right))dv_i \tag{A.2}$$

— Case (b):

$$dy_i = (cz_i + h\left(\frac{1}{n}\sum_{j=1}^{n} z_j\right))dt + \sigma_v dv_i \tag{A.3}$$

for $t \geq 0$ and $1 \leq i \leq n$, where $z_i(t), u_i(t), y_i(t) \in \mathbb{R}$ are the state, the control input and the measured output of the i^{th} agent, respectively. $\{w_i, v_i, 1 \leq i \leq n\}$ denotes a sequence of $2n$ mutually independent standard scalar Wiener processes. The Gaussian initial conditions $z_i(0)$ are mutually independent and are also independent of $\{w_i, v_i, 1 \leq i \leq n\}$. Moreover, σ_v is a positive scalar number, and $b, c, h > 0$.

The problem to be considered is stated as follows.

Problem 1: Design coupled distributed estimation and control strategies based on a feedback control of the form

$$u_i(t) = -f\hat{z}_i(t) \tag{A.4}$$

where $f > 0$ and $\hat{z}_i(t)$ is an estimate of $z_i(t)$, such that each agent's individual cost function given by

$$J_i(u_i) \triangleq \mathbb{E}\int_0^\infty e^{-\rho t}(z_i^2 + ru_i^2)\, dt \tag{A.5}$$

is optimized utilizing only its local information. Here it is assumed that $\rho, r > 0$.

A.4 Coupled Distributed Estimation and Control Algorithm

We combine the Kalman filtering for state estimation and the LQG feedback controller into a closed-loop dynamics model. Noting the information constraints for the agents, the Kalman filtering cannot be directly applied to the n dimensional system. That is also because in our model there is not a central optimizer which can access all other agent's outputs and then form the optimal estimate of the state vector. However, for large n, as in the typical mean field analysis (Huang *et al.* (2007)), we shall assume in the first place that conditions are satisfied so that controlled agents become asymptotically independent (in large population limit), and furthermore the coupling term (mass effect) described by

$$m(t) = h\left(\frac{1}{n}\sum_{j=1}^{n} z_j(t)\right) \tag{A.6}$$

for $t \geq 0$ and $1 \leq i \leq n$, is approximated by a deterministic continuous function $m^*(t)$ defined on $[0, \infty)$ (to be determined later). It is implicitly predicated on the assumption that the coupling in a stochastic process sense, between agent states becomes sufficiently weak as n grows without bounds, and furthermore, that the individual state variance under state estimate feedback remains bounded, so that by the law of large numbers the average in (A.6) converges pointwise a.e. to its (deterministic) mean. This leads to uncoupled measurement equations, and therefore the optimal state estimation for z_i would be given by the standard scalar Kalman filtering. Now in the large but finite population condition, it is expected that the Kalman filtering structure will still produce a satisfactory estimate when $m(t)$ appears in the measurement equations (A.2) and (A.3) but is approximated by $m^*(t)$ when constructing the filtering equation. Here we simply proceed by presuming $m^*(t)$ as a given deterministic function, and the detailed procedure for obtaining this function will be given after the control synthesis is described. In addition, we establish sufficient conditions under which the variance of z_i's remains indeed bounded. This justifies, after the fact, our initial deterministic assumption.

A.4.1 LQG Feedback Controller

Consider only the dynamic model (A.1) (without measurement equation (A.2) or (A.3)) and assume for the moment that the state z_i is completely observable. For minimization of J_i defined by (A.5), the admissible control set is taken as $\mathcal{U}_i = \{u_i | u_i$ is adapted to the σ-algebra $\sigma(z_i(s), s \leq t)$, and $J_i(u_i) < \infty\}$. The set $\mathcal{U}_i$ is nonempty due to controllability of (A.1). Let $f > 0$ be the solution to the algebraic Riccati equation

$$bf^2 + (\rho - 2a)f - \frac{b}{r} = 0 \tag{A.7}$$

Moreover, if one assumes that $\mathbb{E}|z_i(0)|^2 < \infty$ and $\beta_1 = -a + bf > 0$, then the control law

$$u_i(t) = -fz_i(t) \tag{A.8}$$

is stabilizing and further minimizes $J_i(u_i)$ for all $u_i \in \mathcal{U}_i$ (Gelb (1974)).

Assumption 2.

$$a - bf < 0 \tag{A.9}$$

A.4.2 Kalman Filter

Now assume that z_i's are only partially observable and consider either of the two measurement equations (A.2) or (A.3). Approximating and replacing $m(t)$ with the same assumed deterministic function $m^*(t)$ for all agents, the standard (time-varying) Kalman filter would produce the optimal state estimate using the following algorithm (Gelb (1974); Anderson and Moore (1979)),

$$d\hat{z}_i = (a\hat{z}_i + bu_i)dt + K(t)(dy_i - c\hat{z}_i dt) \tag{A.10}$$

$$\frac{dP(t)}{dt} = 2aP(t) + Q - c^2R^{-1}(t)P^2(t) \tag{A.11}$$

$$K(t) = P(t)cR^{-1}(t) \tag{A.12}$$

where the process noise covariance matrice is $Q = \sigma_w^2$ and the measurement noise covariance matrices are chosen as $R_a(t) = (\sigma_v + m^*(t))^2$ and $R_b(t) = \sigma_v^2$ for case (a) and case (b), respectively. Additionally, the separation principle holding, feedback of the state estimate in (A.8) would also produce optimal performance.

Assumption 3. *All agents have mutually independent Gaussian initial conditions with* $\mathbb{E}z_i(0) = \zeta_1$ *and* $\mathbb{E}z_i^2(0) = \zeta_2 > 0$ *for all* i, *where* $\zeta_2 > {\zeta_1}^2$.

Remark A.1. *For agent* i, *the initial condition of the Riccati equation (A.11) is* $Var(z_i(0))$ *which yields the corresponding solution* $P_i(t)$. *Under Assumption 3,* $Var(z_i(0)) = \zeta_2 - {\zeta_1}^2 = \zeta > 0$ *for all* i, *and therefore the same solution* $P(t)$ *is obtained for all agents.*

A.4.3 State Aggregation

Assume $m^*(t) \in C_b[0,\infty)$ is given, where $C_b[0,\infty)$ denotes the set of deterministic, bounded and continuous functions on $[0,\infty)$. For the i^{th} agent, after applying the optimal control law (A.4), the closed loop equation is

$$dz_i = (az_i - bf\hat{z}_i)dt + \sigma_w dw_i \tag{A.13}$$

Denoting $\bar{z}_i(t) = \mathbb{E}z_i(t)$ and taking expectation on both sides of (A.13) gives

$$\frac{d\bar{z}_i}{dt} = a\bar{z}_i - bf\mathbb{E}\hat{z}_i \tag{A.14}$$

with the initial condition $\bar{z}_i|_{t=0} = \mathbb{E}z_i(0)$ assumed and shared by all agents. Also note that in the view of the unbiasedness of the Kalman filter estimate, $\mathbb{E}\hat{z}_i = \mathbb{E}z_i$. Moreover, the population average of means is defined as $\bar{z} \triangleq (1/n)\sum_{i=1}^{n} \bar{z}_i$ and is simply called the population mean. If as assumed $\bar{z}$ becomes deterministic as n goes to infinity, then because of independence of the individual controlled z_i's under Assumption (3), $\bar{z}$ must converge pointwise to its expectation, i.e. to $\mathbb{E}\bar{z}_i$ of a generic agent with initial mean ζ_1, where for the considered uniform population of agents:

$$\frac{d\bar{z}}{dt} = (a - bf)\bar{z} \tag{A.15}$$

which yields

$$\bar{z}(t) = \bar{z}(0)e^{(a-bf)t} \tag{A.16}$$

with condition (A.9) from Assumption 2 guaranteeing boundedness of $\bar{z}(t)$. Here, for simplicity of the analysis, we assume that $\bar{z}(0) \geq 0$.

Furthermore, the population effect $(1/n)\sum_{j=1}^{n} z_j$ is approximated by $\bar{z}$. Since we wish to have

$$m^*(t) \approx m(t) = h\left(\frac{1}{n}\sum_{j=1}^{n} z_j(t)\right) \tag{A.17}$$

for large n, $m^*(t)$ is expressed in terms of the population mean $\bar{z}(t)$ as

$$m^*(t) = h\bar{z}(t) \tag{A.18}$$

Remark A.2. *Under Assumptions 2 and 3, $m^*(t)$ does indeed belong to $C_b[0,\infty)$.*

Remark A.3. *The state aggregation equation (A.15) also holds in the case of perfect observation; that is, it is not affected by the partial observation situation.*

A.4.4 Proposed Algorithm and Closed-loop Dynamics

The coupled distributed estimation and control strategies are presented in Algorithm 1. We proceed by obtaining the resultant closed-loop dynamics of the i^{th} agent and of the population mean.

Solution Algorithm 1

— Initialization

$$\hat{z}_i(0) = \mathbb{E}z_i(0) = \bar{z}(0) \geq 0,\ P(0) = \zeta > 0,\ 1 \leq i \leq n$$

— State Aggregation

$$\bar{z}(t) = \bar{z}(0)e^{(a-bf)t}$$
$$m^*(t) = h\bar{z}(t)$$

— LQG Feedback Controller

$$\rho f = 2af - bf^2 + \frac{b}{r}$$
$$u_i(t) = -f\hat{z}_i(t), \qquad 1 \le i \le n$$

— Kalman filtering

— Case (a):

$$d\hat{z}_i = (a\hat{z}_i + bu_i)dt + K(t)(dy_i - c\hat{z}_i dt)$$
$$\frac{dP(t)}{dt} = 2aP(t) + \sigma_w^2 - c^2(\sigma_v + m^*)^{-2}P^2(t)$$
$$K(t) = P(t)c(\sigma_v + m^*)^{-2}$$

— Case (b):

$$d\hat{z}_i = (a\hat{z}_i + bu_i)dt + K(t)(dy_i - (c\hat{z}_i + m^*)dt)$$
$$\frac{dP(t)}{dt} = 2aP(t) + \sigma_w^2 - c^2\sigma_v^{-2}P^2(t)$$
$$K(t) = P(t)c\sigma_v^{-2}$$

Case (a):

Defining the estimation error as

$$\tilde{z}_i = z_i - \hat{z}_i \tag{A.19}$$

and replacing (A.2) and (A.4) in (A.10), yields

$$d\hat{z}_i = (a - bf)\hat{z}_i dt + P(t)c^2(\sigma_v + m^*)^{-2}\tilde{z}_i dt + P(t)c(\sigma_v + m^*)^{-2}(\sigma_v + h\left(\frac{1}{n}\sum_{j=1}^{n} z_j\right))dv_i \tag{A.20}$$

Subsequently, using (A.1), (A.19) and (A.20) it follows that

$$dz_i = (a - bf)z_i dt + bf\tilde{z}_i dt + \sigma_w dw_i \tag{A.21}$$

and

$$d\tilde{z}_i = (a - P(t)c^2(\sigma_v + m^*)^{-2})\tilde{z}_i dt + \sigma_w dw_i - P(t)c(\sigma_v + m^*)^{-2}(\sigma_v + h\left(\frac{1}{n}\sum_{j=1}^{n} z_j\right))dv_i \tag{A.22}$$

In addition, letting $z'_n = (1/n)\sum_{i=1}^{n} z_i$, $\hat{z}'_n = (1/n)\sum_{i=1}^{n} \hat{z}_i$, $\tilde{z}'_n = (1/n)\sum_{i=1}^{n} \tilde{z}_i$, $w' = (1/\sqrt{n})$ $\sum_{i=1}^{n} w_i$, and $v' = (1/\sqrt{n})\sum_{i=1}^{n} v_i$, where w' and v' are two independent standard Wiener processes, we get

$$dz'_n = (a - bf)z'_n dt + bf\tilde{z}'_n dt + \frac{1}{\sqrt{n}}\sigma_w dw' \tag{A.23}$$

and

$$d\tilde{z}'_n = (a - P(t)c^2(\sigma_v + m^*)^{-2})\tilde{z}'_n dt + \frac{1}{\sqrt{n}}\sigma_w dw' - \frac{1}{\sqrt{n}}P(t)c(\sigma_v + m^*)^{-2}(\sigma_v + hz'_n)dv' \tag{A.24}$$

Case (b):

Similarly, using (A.3) in place of (A.2), it follows that

$$dz_i = (a - bf)z_i dt + bf\tilde{z}_i dt + \sigma_w dw_i, \tag{A.25}$$

$$\begin{aligned} d\tilde{z}_i =& (a - P(t)c^2\sigma_v^{-2})\tilde{z}_i dt - P(t)c\sigma_v^{-2}h\left(\frac{1}{n}\sum_{j=1}^{n} z_j\right)dt + P(t)c\sigma_v^{-2}m^*(t)dt \\ &+ \sigma_w dw_i - P(t)c\sigma_v^{-1}dv_i \end{aligned} \tag{A.26}$$

and

$$dz'_n = (a - bf)z'_n dt + bf\tilde{z}'_n dt + \frac{1}{\sqrt{n}}\sigma_w dw', \tag{A.27}$$

$$\begin{aligned} d\tilde{z}'_n =& (a - P(t)c^2\sigma_v^{-2})\tilde{z}'_n dt - P(t)c\sigma_v^{-2}hz'_n dt + P(t)c\sigma_v^{-2}m^*(t)dt \\ &+ \frac{1}{\sqrt{n}}\sigma_w dw' - \frac{1}{\sqrt{n}}P(t)c\sigma_v^{-1}dv' \end{aligned} \tag{A.28}$$

A.5 Stability Analysis

For stability analysis of the closed-loop dynamics (A.23)-(A.24) and (A.27)-(A.28) we make use of the following concepts of boundedness for solutions of stochastic differential equations.

Definition A.1. *(Reif et al. (2000)) The stochastic process $x(t)$ is said to be stochastically sample path bounded, if for every $\delta > 0$ there is a $\beta(\delta) > 0$ such that*

$$\mathbb{P}[\sup_{t\geq 0} \|x(t)\| \leq \beta(\delta)] \geq 1 - \delta \tag{A.29}$$

Definition A.2. *(Reif et al. (2000)) The stochastic process $x(t)$ is said to be exponentially bounded in mean square, if there are real numbers $\eta, \nu, \varphi > 0$ such that*

$$\mathbb{E}\, \|x(t)\|^2 \leq \varphi\, \|x(0)\|^2 \exp(-\eta t) + \nu \tag{A.30}$$

holds for every $t \geq 0$.

Definition A.3. *(Reif et al. (2000); Mao and Yuan (2006)) Consider the continuous-time stochastic process described by the Itô stochastic differential equation,*

$$dx(t) = F(x(t), t)dt + G(x(t), t)d\bar{w}(t) \tag{A.31}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the state, and $\bar{w}(t) \in \mathbb{R}^{n_w}$ is a standard Wiener process. Moreover, the nonlinear functions F and G are assumed to be continuously differentiable, and such that (A.31) has a unique solution. Consider a nonnegative function $V(x(t), t)$ which is continuously twice differentiable in x and once in t, i.e., $V(x, t) \in C^{2,1}$. The differential generator of (A.31) associated with the Lyapunov function $V(x, t)$ is defined by

$$\mathcal{L}V(x,t) = \frac{\partial V}{\partial t}(x,t) + \frac{\partial V}{\partial x}(x,t)F(x,t) + \frac{1}{2}\sum_{i=1}^{n_x}\sum_{j=1}^{n_x} \frac{\partial^2 V}{\partial x_i \partial x_j}(x,t)[G(x,t)G^T(x,t)]_{i,j} \tag{A.32}$$

where $x = [x_1, \ldots, x_{n_x}]^T$, and $[G(x,t)G^T(x,t)]_{i,j}$ is the matrix element of $G(x,t)G^T(x,t)$ in the i^{th} row and the j^{th} column. Furthermore, the sum in (A.32) can be simplified as

$$\begin{aligned}\sum_{i=1}^{n_x}\sum_{j=1}^{n_x} \frac{\partial^2 V}{\partial x_i \partial x_j}(x,t)[G(x,t)G^T(x,t)]_{i,j} &= tr(G(x,t)G^T(x,t)Hess_x[V(x,t)]) \\ &= tr(Hess_x[V(x,t)]G(x,t)G^T(x,t)) \end{aligned} \tag{A.33}$$

where $Hess_x[\cdot]$ denotes the Hessian matrix with respect to x as the variable vector.

Lemma A.1. *(Reif et al. (2000)) Assume there is a sufficiently smooth function* $V(x(t),t) \in C^{2,1}$ *of the stochastic process* $x(t)$ *in (A.31) and real numbers* $\vartheta_{\min}, \vartheta_{\max}, \mu, \alpha > 0$ *such that*

$$\vartheta_{\min} \|x(t)\|^2 \le V(x(t),t) \le \vartheta_{max} \|x(t)\|^2 \tag{A.34}$$

and

$$\mathcal{L}V(x(t),t) \le -\alpha V(x(t),t) + \mu \tag{A.35}$$

are satisfied. Then the stochastic process $x(t)$ *is exponentially bounded in mean square, i.e.,*

$$\mathbb{E}\|x(t)\|^2 \le \frac{\vartheta_{\max}}{\vartheta_{\min}} \|x(0)\|^2 \exp(-\alpha t) + \frac{\mu}{\vartheta_{\min}\alpha} \tag{A.36}$$

for every $t \ge 0$*. Moreover, the stochastic process is sample-path bounded.*

Lemma A.2. *Each system described by scalar dynamics (A.1) and either of*

$$dy_i = cz_i dt + (\sigma_v + m^*(t))dv_i \tag{A.37}$$

or

$$dy_i = (cz_i + m^*(t))dt + \sigma_v dv_i \tag{A.38}$$

is uniformly detectable.

Proof: For $c > 0$, any real number Λ such that $\Lambda < \frac{-a}{c}$ yields $a + \Lambda c < 0$. Therefore, according to definition 4.1 in Reif *et al.* (2000), the system is uniformly detectable. □

Lemma A.3. *(Reif et al. (2000)) For each uniformly detectable system described by (A.1) and either of (A.37) or (A.38), there are real numbers* $p_{\min}, p_{max} > 0$ *such that the solution* $P(t)$ *of the scalar Riccati differential equation (A.11) satisfies the bounds*

$$p_{\min} \le P(t) \le p_{max} \tag{A.39}$$

for every $t \ge 0$*.*

Theorem A.1. *If*

$$f > \frac{2a}{b} \tag{A.40}$$

and also if there exists a fixed real number l *such that*

$$\frac{2p_{\max}c^2h^2}{n\sigma_v^4(bf-2a)} < l < \frac{\sigma_w^2}{bfp_{\max}^2} + \frac{c^2}{bf(\sigma_v + h\bar{z}(0))^2}\,, \tag{A.41}$$

then the stochastic process $x(t) = \begin{bmatrix} z'_n(t) \\ \tilde{z}'_n(t) \end{bmatrix}$ *verifying the closed-loop dynamics (A.23)-(A.24) for* $1 \leq i \leq n$*, is exponentially bounded in mean square and stochastically sample-path bounded. Specifically,*

$$\mathbb{E}({z'_n}^2(t) + {\tilde{z}'_n}^2(t)) \leq \frac{\max\{l, \frac{1}{p_{\min}}\}}{\min\{l, \frac{1}{p_{\max}}\}}({z'_n}^2(0) + {\tilde{z}'_n}^2(0))e^{-\alpha t} + \frac{(l + \frac{1}{p_{\min}})\frac{\sigma_w^2}{n} + \frac{2p_{\max}c^2}{n\sigma_v^2}}{\min\{l, \frac{1}{p_{\max}}\}\alpha} \tag{A.42}$$

holds with

$$\alpha = \frac{1}{\max\{l, \frac{1}{p_{\min}}\}} \min\{l(bf - 2a) - \frac{2p_{\max}c^2h^2}{n\sigma_v^4}, -lbf + \frac{\sigma_w^2}{p_{\max}^2} + \frac{c^2}{(\sigma_v + h\bar{z}(0))^2}\} \tag{A.43}$$

Proof: Choosing

$$V(x(t), t) = x^T(t)\Pi(t)x(t) \tag{A.44}$$

with

$$\Pi(t) = \begin{bmatrix} l & 0 \\ 0 & P^{-1}(t) \end{bmatrix} \tag{A.45}$$

where l is a fixed real number verifying (A.41) and applying Lemma (A.3) we can write

$$l{z'_n}^2 + \frac{1}{p_{\max}}{\tilde{z}'_n}^2 \leq V(x, t) \leq l{z'_n}^2 + \frac{1}{p_{\min}}{\tilde{z}'_n}^2 \tag{A.46}$$

Therefore, (A.34) is verified with $\vartheta_{\min} = \min\{l, \frac{1}{p_{\max}}\}$ and $\vartheta_{\max} = \max\{l, \frac{1}{p_{\min}}\}$. □

Next, considering the dynamic equations (A.23)-(A.24) as in the form of (A.31) with

$$F(x, t) = \begin{bmatrix} (a - bf)z'_n + bf\tilde{z}'_n \\ (a - P(t)c^2(\sigma_v + m^*)^{-2})\tilde{z}'_n \end{bmatrix}, \bar{w}(t) = \begin{bmatrix} w' \\ v' \end{bmatrix} \tag{A.47}$$

and

$$G(x, t) = \begin{bmatrix} \frac{1}{\sqrt{n}}\sigma_w & 0 \\ \frac{1}{\sqrt{n}}\sigma_w & -\frac{1}{\sqrt{n}}P(t)c(\sigma_v + m^*)^{-2}(\sigma_v + hz'_n) \end{bmatrix} \tag{A.48}$$

and using (A.32) we have

$$\mathcal{L}V(x,t) = x^T(t)\frac{d\Pi(t)}{dt}x(t) + 2x^T(t)\Pi(t)F(x,t) + \frac{1}{2}tr[G(x,t)G^T(x,t)Hess_x(V(x,t))] \tag{A.49}$$

which can be expressed as

$$\begin{aligned}\mathcal{L}V =& (2aP^{-1}(t) - 2c^2(\sigma_v + m^*)^{-2} - \dot{P}(t)P^{-2}(t))\tilde{z}_n'^{\,2} + 2l(a-bf)z_n'^{\,2} + 2lbfz_n'\tilde{z}_n' \\ &+ tr[GG^T\Pi(t)]\end{aligned} \tag{A.50}$$

Expressing $\dot{P}(t)$ by means of Riccati differential equation (A.11), and applying the inequality $2z_n'\tilde{z}_n' \le z_n'^{\,2} + \tilde{z}_n'^{\,2}$ yields

$$\begin{aligned}\mathcal{L}V(x,t) \le& l(2a-bf)z_n'^{\,2} + (lbf - \sigma_w^2P^{-2}(t))\tilde{z}_n'^{\,2} - c^2(\sigma_v + m^*)^{-2}\tilde{z}_n'^{\,2} + \frac{1}{n}(l + P^{-1}(t))\sigma_w^2 \\ &+ \frac{1}{n}P(t)c^2(\sigma_v + m^*)^{-4}(\sigma_v + hz_n')^2\end{aligned} \tag{A.51}$$

In addition, using the bounds from Lemma A.3, applying the inequality $2\sigma_v hz_n' \le \sigma_v^2 + h^2 z_n'^{\,2}$ and noting that $\min_t\{\sigma_v + m^*(t)\} \ge \sigma_v > 0$ and also $\max_t\{\sigma_v + m^*(t)\} \le \sigma_v + h\bar{z}(0)$, after some simplification in the right-hand side of (A.51), it can be written as

$$\begin{aligned}\mathcal{L}V \le& (l(2a-bf) + \frac{2p_{\max}c^2h^2}{n\sigma_v^4})z_n'^{\,2} + (l + \frac{1}{p_{\min}})\frac{\sigma_w^2}{n} + \frac{2p_{\max}c^2}{n\sigma_v^2} \\ &+ (lbf - \sigma_w^2P^{-2}(t) - c^2(\sigma_v + h\bar{z}(0))^{-2})\tilde{z}_n'^{\,2}\end{aligned} \tag{A.52}$$

Enforcing conditions (A.40) and (A.41) in (A.52) and using the bounds from Lemma A.3, inequality (A.35) holds with α defined in (A.43) and

$$\mu = (l + \frac{1}{p_{\min}})\frac{\sigma_w^2}{n} + \frac{2p_{\max}c^2}{n\sigma_v^2} \tag{A.53}$$

Therefore, it follows that (A.42) holds according to (A.36). This ends the proof for case (a).□

Theorem A.2. *If*

$$f > \frac{a}{b}, \tag{A.54}$$

then the stochastic process $x(t) = \begin{bmatrix} z_n'(t) \\ \tilde{z}_n'(t)\end{bmatrix}$ *verifying the closed-loop dynamics (A.27)-(A.28) for* $1 \le i \le n$*, is exponentially bounded in mean square and stochastically sample-path* *bounded. Specifically,*

$$\mathbb{E}({z'_n}^2(t) + {\tilde{z}'_n}^2(t)) \leq \frac{\max\{l, \frac{1}{p_{\min}}\}}{\min\{l, \frac{1}{p_{\max}}\}}({z'_n}^2(0) + {\tilde{z}'_n}^2(0))e^{-\alpha t} + \frac{(l + \frac{1}{p_{\min}})\frac{\sigma_w^2}{n} + \frac{p_{\max}c^2}{n\sigma_v^2} + \sigma_v^{-2}h^2\bar{z}^2(0)}{\min\{l, \frac{1}{p_{\max}}\}\alpha} \tag{A.55}$$

holds with

$$\alpha = \frac{1}{\max\{l, \frac{1}{p_{\min}}\}}\min\{2l(bf - a), \frac{\sigma_w^2}{p_{\max}^2}\} \tag{A.56}$$

and a fixed real number $l = \frac{ch\sigma_v^{-2}}{bf}$ *.*

Proof: Considering the dynamic equations (A.27)-(A.28) in place of equations (A.23)-(A.24) and following a similar procedure as in Theorem A.1, this theorem can be proved for case (b). The details are omitted here for brevity. □

Remark A.4. *Note that under our stability condition (A.40), inequality (A.42) is consistent with our initial assumption that as* n *goes to infinity* $z'_n(t)$ *becomes deterministic. Indeed, it confirms that as* n *goes to infinity, the variance of* $z'_n(t)$ *under measurement model (a), goes to zero, provided the initial mean of* z'_n *is known to all agents. Similarly, under stability condition (A.54), inequality (A.55) indicates that if the initial agents' mean estimate of* $\bar{z}(0)$ *is correct and equal to zero, the variance of* $z'_n(t)$ *under measurement model (b) , goes to zero. The same result can be shown to remain true if the initial mean of* $\bar{z}(0)$ *is strictly positive, but again correctly estimated by all agents.*

A.6 Numerical Example

The following numerical values were used with noise model (a): $a = -0.5$, $b = c = h = \sigma_w = \sigma_v = \rho = r = 1$, $\mathbb{E}z_i(0) = 10$ and $Var(z_i(0)) = 1$. The simulation results as depicted in Figs. A.1 and A.2 illustrate a case where f^*_{LQG} satisfies the bound of Theorem 1, thus guaranteeing convergence of the mean state when f^*_{LQG} is used, to a deterministic value. In particular, Fig. A.1 illustrates the LQ cost J as a function of f, where the black vertical line indicates f^*_{LQG}, the dashed green vertical line represents the bound $\frac{2a}{b}$, and the crosses denote the points where one of the two conditions of Theorem 1 is not met. In this case, f^*_{LQG} induces a Nash equilibrium. It also happens to be a socially optimal equilibrium. However, in general the separation principle will not hold for noise model (a) even as n goes to infinity.

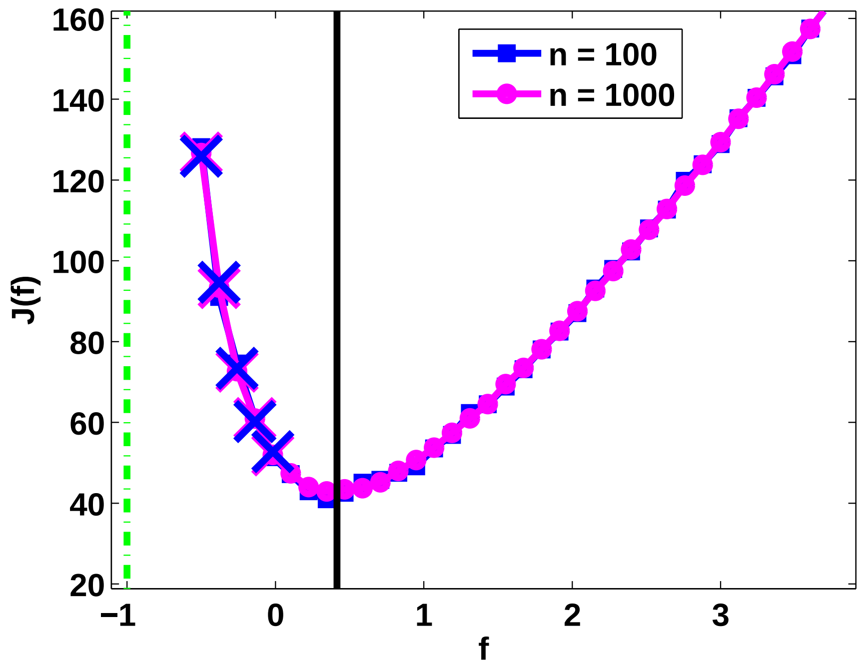

Figure A.1 The LQ cost J as a function of f

A.7 Conclusion

This paper addressed the distributed decision-making in a system of uniform agents coupled via their measurement equations, whereby each agent has noisy measurements of its own state. Specifically, a distributed estimation and control algorithm was developed using a decentralized control synthesis in which each agent utilizes an estimate based on its local information and a priori (shared) information on the initial mean state estimate for its control strategy. One special feature of the proposed algorithm is the fact that it combines the Kalman filtering for state estimation and the linear-quadratic-Gaussian (LQG) feedback controller based on the anticipation of the collective effect (mean field) of all agents and using the state aggregation technique to anticipate that effect. It was proved that under certain conditions the closed-loop dynamics is exponentially bounded in mean square and stochastically sample-path bounded.

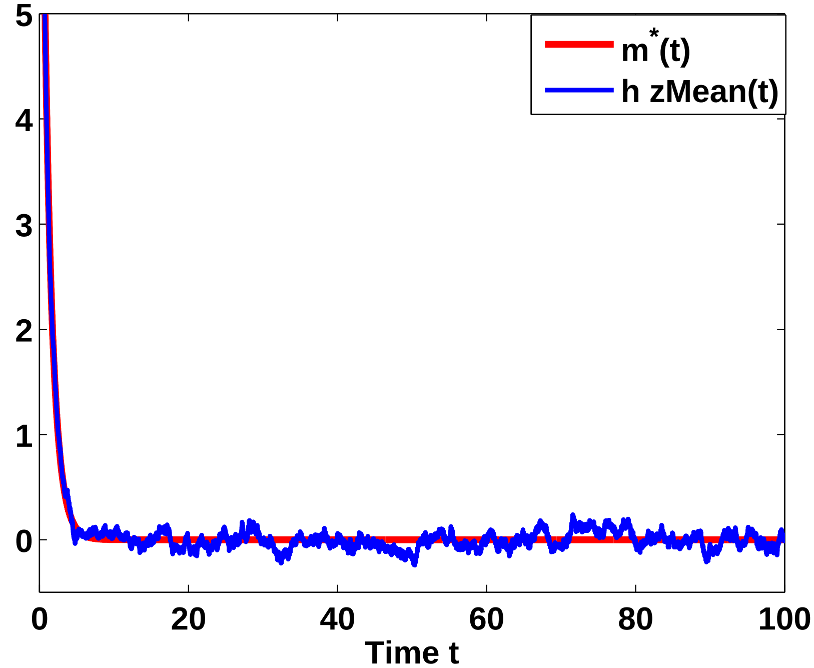

Figure A.2 Comparison between $m^*(t)$ and $h\left(\frac{1}{n}\sum_{j=1}^{n} z_j(t)\right)$ for $n = 100$